Sports Core Literacy Education of College Students under the Background of "Internet+" and Big Data

Xian Liu¹,a, XiangPing Mei²,b and JianQiang Guo¹,c*

¹School of Physical Education, Changzhou University, Changzhou 213164, Jiangsu, China
²Huaide College, Changzhou University, JingJiang 214500, Jiangsu, China

¹liuxian@cczu.edu.cn, bmxp@sina.com, cgjq@cczu.edu.cn

*Corresponding author

Keywords: "Internet +" and Big Data, Sports Core Literacy, Cluster Analysis, Generalization Ability

Abstract: Since the new century, people have become more and more aware of the importance of sports, but the current teaching methods of physical education in colleges and universities are relatively rough, and it is difficult to achieve the standard for core physical education. In order to solve the existing problems of sports core literacy education and optimize the resource allocation of college courses, this article uses cluster analysis to classify some university students in this city under the background of "Internet +" and big data, and provides certain information for other researchers. Through the empirical analysis of the core sports literacy evaluation index system in the education stage, the K-means and GA indicators in the cluster analysis method further prove the scientificity and rationality of sports literacy. The experimental results show that through clustering and then the core literacy education for students, it can solve the contradiction between the unevenness of students in physical education and the synchronization of teaching requirements. Students' physical performance is about 30% higher than traditional teaching methods. The awareness of cooperation between students reaches 0.8 awareness, solves the contradiction between popularization and improvement, and solves the problems of polarization and transformation of underachievers, so that their personalities can be fully displayed and developed, and the core literacy education of physical education in ordinary colleges and universities will be on the track of quality education. This shows that in the context of "Internet +" and big data, cluster analysis can play an extremely important role in college physical education and teaching.

1. Introduction

After the 1990s, some foreign countries began to build their own basic framework of student core literacy. Our country is influenced by the international community and at the same time aware of the importance of core literacy. According to our country’s development status and people’s
living standard at that time, our country’s core literacy has been constructed, and the core literacy of students of different ages has been divided into stages [1]. In recent years, my country has been advocating quality education, and physical education is one of the subjects most easily overlooked in exam-oriented education. The cultivation of physical literacy has long-term significance in the future. Modern physical education has placed core physical literacy at the center. This is particularly important. It is conducive to improving the physical, psychological, and other personal qualities of students. Compared with the previous ordinary physical education, it is very different for students. It is not only necessary to carry out physical education knowledge and improve physical quality, but also to help students establish good beliefs, have healthy values, cultivate students' hands-on skills, communication skills, and participate in social practice. Physical education is a comprehensive subject. The compulsory education stage should not only cultivate students' physical fitness, but also take the cultivation of core physical literacy as the focus of physical education [2].

Quality education has been implemented in our country for many years, but there is no obvious immediate effect, and core physical literacy is the most neglected in education. It is well known that physical education can not only exercise the body, but also benefit students’ intellectual development and increase concentration. Have more energy, so the core sports literacy research is very important [3]. In the face of modernized higher education, with human modernization as the core, we are committed to cultivating creators who adapt to the future and change the future. How to tap the high-level qualities that have been condensed in human beings with the most core competitiveness has become the focus of various countries and international organizations. Therefore, from the perspective of higher education modernization, exploring the cultivation and development of college students’ core literacy of sports is not only the inheritance and innovation at the current stage based on the implementation of quality education, but also the urgent need for my country to improve the quality of talent training in colleges and universities under the background of double first-class construction. The theoretical and practical problems solved [4].

Overseas research on sports core literacy education is relatively early. Mr. Huang believes that outreach training can be added to college physical education curriculum to form a new physical education curriculum content to strengthen students’ quality training [5]; Xu Jinguang believes that college physical education has problems. Because it is not combined with lifelong sports, he believes that the advantages of important sports should be absorbed to solve the problem of college sports [6]; Peng Qingjun believes that the most important problem of college physical education is that the relevant resources are not fully developed and utilized, so he It has studied how to maximize the development and utilization of existing resources in college physical education courses [7]. In China, related research started late. Xin Tao pointed out that core literacy is not based on a single aspect, but from a comprehensive perspective. Students should be able to adapt to social development in the future society, and should For all-round development, these abilities require students to learn on the basis of the school’s mastery of basic literacy. Education is vital to students. Schools should conduct a deeper analysis of education and address "what kind of people to cultivate". Conduct education [8]. Ma Hongxia investigated Henan college students' sports related aspects through five aspects: sports knowledge dimension, sports awareness dimension, sports morality dimension, sports behavior dimension and factors that affect the improvement of students' sports literacy [9]. Qiu Jianguo conducted a survey of youth sports literacy through youth sports knowledge, sports ability, sports awareness, sports morality, sports personality, and sports behaviors. He pointed out some problems in my country: there are serious regional differences in basic physical education, and student sports and health are related. Knowledge is lacking, the
overall level of physical fitness is not high, and physical literacy is in a state of "uncoordinated" [10]. These studies have a certain reference effect for this article, but because the experimental conclusions are not completely reasonable, the research samples are small and the conclusions are not universal enough to be convincing.

By summarizing and analyzing previous scholars’ research, this research has recognized the deficiencies in their articles and improved the research methods. It not only uses expert argumentation to establish a core sports literacy index system, but also uses professional statistical methods to conduct objectively. The establishment of the index system uses the clustering method to divide the constituent elements into sports knowledge, sports awareness, sports behavior, sports personality, and sports morality. It also introduces the index of physical fitness level to improve the composition system of sports literacy. And in the analysis part of the article, the independent sample T test and one-way analysis of variance were used to analyze the differences in core physical literacy among students of different genders and different grades, making the research on the core literacy indicator system more valuable and more valuable. It can deeply analyze the differences between them. Relevant departments, schools and society can teach these students in accordance with their aptitude, which part is lacking and which part, so that the teaching is targeted and not blindly teaching.

2. College Students’ Sports Core Literacy Education Methods

2.1. Definition of the Core Literacy of College Students

Academia often uses the term "quality" as a description of the development status and development results of college students after a complete education. Quality condenses people's knowledge, abilities, and inner psychological qualities, and the relationship between literacy and knowledge, abilities, and qualities has become a definition the core literacy of college students must be studied. In view of the fact that the core literacy of college students itself is a combined concept, it incorporates the theoretical connotations of "literacy" and "core literacy", clarifies the relationship between literacy and basic concepts such as knowledge and ability in the traditional educational theoretical paradigm, and combines the characteristics of their own times. Discuss "literacy" and "core literacy" at the academic level [11].

"Quality" is a word derived from tradition and modernity, clarifying its relationship with knowledge, ability and quality in traditional Chinese education will help to better understand its tradition and modernity, and facilitate higher education practice. Literacy and knowledge: Firstly, knowledge is no longer an "objective truth" or "fixed fact" but has become a research object for the future development of mankind, supporting the resources that mankind can use to solve practical problems; secondly, it is to promote knowledge into concepts, knowledge in the information age The rate of attenuation and update of the knowledge is unprecedentedly accelerated, but the concepts or thoughts embodied in knowledge are relatively stable; again, the students’ personal knowledge becomes the thinking and enlightenment generated by the individual accepting subject knowledge, digesting and understanding, and solving objective problems in the life world. Building on life understanding and experience is the basis and prerequisite for forming literacy. Therefore, the key to developing student literacy is to respect students’ personal knowledge. Finally, all knowledge can only become the object of students’ exploration and practice in the learning process. Literacy development process [12-13].

Literacy and ability: First of all, ability is not fixed, it will continue to develop and improve with the changes of the times, and literacy has a long "shelf life". Secondly, the development of literacy
is closely integrated with inquiry and practice, which directly refers to the solution of practical problems, while the "ability" in the traditional sense is slightly weak. Literacy and quality: First of all, quality focuses on innate inheritance, literacy focuses on acquired cultivation, quality represents a person's innate physiological and anatomical characteristics, literacy is the preservation of learning and internalization; secondly, although quality and literacy can be obtained through school education, but literacy focuses more on "education." Finally, the quality inherent in people is a static existence, and the literacy bred by context reflects dynamics [14-15].

"From the perspective of building blocks, it includes the integration and integration of key abilities, necessary qualities, and values. Key abilities refer to the important skills that students use to analyze and solve problems by using existing objective conditions in their future development. The foundation is the cornerstone of a smart life, which contains people's creativity, initiative and inward development; the essential character shows the important qualities that students have experienced after the completion of higher education and the cultivation of diverse knowledge condensed on people, and they are human. The foundation is the cornerstone of a moral life. It contains human morality, spirituality and altruism; value concept refers to the conceptual representation of the student's value choice with what purpose and value standard, key abilities, and essentials Character and values are unified in the core literacy of college students, and the three are inseparable [16].

2.2. "Internet +" and Big Data

The development of the big data industry is closely related to big data and its use. Although it originated from industry practice, academic research on the "big data industry" lags far behind the development of practice [17-18]. From a domestic perspective, the current research on the big data industry is mainly based on government industrial policies and plans, industrial development suggestions, comparisons of domestic and foreign big data industries, and industrial development influencing factors. There is a lack of appropriate theoretical perspectives on the internal big data industry. Research on constituent elements and governance mechanisms; from a foreign perspective, although there are not many related studies, scholars have begun to discuss the big data industry from the perspective of business ecology. Big data is generally achieved through the following methods.

\[ m_u = \sum_{b=1}^{n} \lambda_{ab} g_{ab} \]  \hspace{1cm} (1)

Among them \( m_{ab} \) is the individual's contribution to the overall degree of ordering, and \( \lambda_{ab} \) is the weight of each order parameter. Therefore.

\[ y(kT + t_i) = \frac{1}{\alpha(z)} \sum_{j=1}^{r} \beta_j(z)\mu(kT + t_{j-1}) + v(kT + t_i) \]  \hspace{1cm} (2)

Can be transformed into:

\[ \alpha(z) = 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + ... + \alpha_n z^{-n} \]  \hspace{1cm} (3)

\[ \beta_j(z) = \beta_j^0 + \beta_j^1 z^{-1} + \beta_j^2 z^{-2} + ... + \beta_j^n z^{-n} \]  \hspace{1cm} (4)

Its function \( s(kT + t_{i-1}) \) is to move the sampling signal \( s(kT + t_{i-1}) \) in time.
backward by 1 non-uniform sampling interval, and a new transfer function model is proposed:

\[ y(kT + t) = \frac{B_i(\delta)}{A_i(\delta)} \overline{r}(kT + t) + v(kT + t) \]  

(5)

The big data ecology is divided into three levels, namely the core value chain at the micro level, the extended value chain at the meso level, and the big data ecology at the macro level [19]. Among them, the core value chain is centered on the data value chain, including direct data suppliers and data value distribution channels; the extended value chain is centered on the core value chain, consisting of providers, data markets, data suppliers’ suppliers, complementary data products and Service providers, direct data end users, etc.; and the macro-level big data ecology mainly refers to some related organizations in the periphery, such as government agencies, regulatory agencies, investors, venture capital & incubators, industry associations, academics and research Institutions, standardization organizations, start-ups and entrepreneur groups, as well as various other competitors, stakeholders, peripheral members, etc. The recommendation algorithm through big data is shown in Figure 1:

![Figure1: Recommendation algorithm composition](image)

For the big data ecology, the diversity of members is crucial [20]. Diversity is an ecological concept. Various organisms in the ecology play different important roles in the environment. Many complete food chains and complex food webs have been formed between species and species, and between organisms and the environment. The circle constitutes a virtuous circle of material and energy flow. Once the food chain is broken, the function will not be able to perform normally. Similar to the natural ecology, diversity is also indispensable to the big data business ecology: first, the diversity of members plays a buffer role for its response to environmental uncertainty; second, the value creation of diversity on the big data business ecology it is of great benefit. For example, in order to build a data-centric business ecosystem, Alibaba has successively invested or acquired many Internet companies with a large amount of high-quality data, such as Sina Weibo and Didi Chuxing. This has played a huge role in the creation of ecological value. Role; third, diversity is a prerequisite for self-organization of the big data business ecology [21].

2.3. Cluster Analysis Method

Cluster analysis usually refers to the process of decomposing a group of data sets with a large number of data objects into many clusters, and in the division process, we need to divide according
to the similarity or certain distance between each data object. The high similarity is classified into the same cluster, and the large difference or low similarity is classified into different clusters [22-23]. Clustering is to cluster all records into different clusters when the number of classes in the database is unknown. According to different classification standards, clustering analysis can be divided into a variety of algorithms. Here we briefly introduce several common clustering algorithms. They are hierarchical method, partition method, clustering algorithm based on density and grid. The characteristics of cluster analysis are simple calculations, and the results presented are relatively intuitive and easy to understand. In the teaching field, cluster analysis can analyze student learning and inspect teaching quality [24]. The algorithm flow of the clustering analysis method is shown in Figure 2:

![Flow chart of clustering analysis algorithm](attachment:algorithm_flow_chart.png)

With the advent of the era of big data, data structures have become more and more complex, making it difficult to accurately find the optimal number of clusters due to their own shortcomings in some existing effectiveness indicators [25]. In response to this situation, this section evaluates the effectiveness of clustering through generalization ability, and designs a new clustering effectiveness index—generalization ability (GA) indicator, which can measure the generalization of the current clustering results. Evaluation of chemical ability to judge the pros and cons of the clustering results.

Let be the initial sample space $X = \{x_1, x_2, \ldots, x_n\}$, randomly split the sample space $X$ according to a certain ratio, split into the training sample space $x_{kr}$ and the test sample space $x_{kr}$, cluster the training sample space $x_{kr}$ and the test sample space $x_{kr}$, and let the training $sT_{kr} = \{x_{kr}, r_{kr}\}$ and cluster
space, $T_{kr} = \{x_{kr}, r_{kr}\}$ is the clustering space of the test set, where:

\[
x_{kr} = \{x_{kr}^1, x_{kr}^2, \ldots, x_{kr}^n\}, \quad r_{kr} = \{r_{kr}^1, r_{kr}^2, \ldots, r_{kr}^m\}
\]

(6)

\[
x_{ke} = \{x_{ke}^1, x_{ke}^2, \ldots, x_{ke}^n\}, \quad r_{ke} = \{r_{ke}^1, r_{ke}^2, \ldots, r_{ke}^m\}
\]

(7)

Function $M$ is the best classification function obtained by training the machine learning classification algorithm on the clustering space of the training set. Function $M$ is the mapping rule, which satisfies

\[
r = M(x)
\]

(8)

Make arbitrary

\[
x_{kr}^i \in X_{kr}
\]

(9)

Yes and only

\[
r_{kr}^j \in R_{kr}
\]

(10)

Make

\[
r_{kr}^j = M(x_{kr}^i)
\]

(11)

Among them

\[
x_{kr} = \{x_{kr}^1, x_{kr}^2, \ldots, x_{kr}^n\}
\]

(12)

\[
Y_{kr} = M(X_{kr}) = \{y_{kr}^1, y_{kr}^2, \ldots, y_{kr}^n\}
\]

(13)

The number of correctly classified sample points is $nt$, namely

\[
nt = \sum_{i=1}^{n} f(r_{kr}^i, y_{kr}^i)
\]

(14)

Evaluating the effectiveness of clustering through generalization ability allows us to analyze the students’ comprehensive test scores and discover the students’ learning in a certain course. According to the characteristics of this course, combined with the characteristics of the students’ own existence, the effect on the students’ learning Make a reasonable and scientific evaluation, provide certain guidance for teaching activities, and the final results will inspire teachers' teaching methods.

The proposed GA index is proposed to evaluate the validity of the clustering results. The GA indicator is based on the current clustering results and evaluates the clustering results from the generalization ability in guided learning, that is, it is considered that the pros and cons of the clustering results are related to the generalization ability of predicting unknown samples, so it is different from the existing The clustering effectiveness index, whether it is an external effectiveness indicator or an internal effectiveness indicator. The GA indicator splits the original data set into a training set and a data set, and clusters them separately. Use the clustering results of the training set to perform machine learning to construct a classifier, and use this classifier to predict the test set, and then compare the prediction results and the clustering results. Therefore, the GA indicator is similar to the external effectiveness indicator. The difference is that the commonly used external effectiveness indicator is difficult to obtain the true category of the data set. The GA indicator
constructs a classifier and replaces the real category of the test data set with the result of the classifier prediction. Solve the problem that the real category of the data set is difficult to obtain.

3. College Students' Sports Core Literacy Education Experiment

3.1. Subjects

This article makes full use of the research results in the field of physical education and teaching in colleges and universities to achieve the goal of cultivating the core literacy of students in physical education in colleges and universities. Through in-depth research on the existing problems in colleges and universities, a system for optimizing college physical education is constructed. Physical quality statistics, add up all students' usual scores and final exam scores to get a comprehensive score.

3.2. Data Preprocessing

The data we collected in this experiment makes the students’ comprehensive scores almost non-noisy. The work we need to do with these data is just to fill in the final scores due to the fact that the students cannot take the exam in time, then we can use Fill in the vacancy data manually by filling in the vacancy and the average value. K-means clustering is performed on the test set and the training set, the selection of k value adopts the exhaustive method, and k=2, 3, 4, 5 are selected respectively, and the clustering is performed in turn. After that, the test set is classified according to the clustering results of the training set, and the obtained classification results are compared with the previous clustering results, and the GA indices under different k values are calculated. The results are shown in Table 1:

<table>
<thead>
<tr>
<th>K</th>
<th>N</th>
<th>Nke</th>
<th>GA</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>175</td>
<td>265</td>
<td>0.664</td>
</tr>
<tr>
<td>3</td>
<td>263</td>
<td>265</td>
<td>0.989</td>
</tr>
<tr>
<td>4</td>
<td>227</td>
<td>265</td>
<td>0.858</td>
</tr>
<tr>
<td>5</td>
<td>235</td>
<td>265</td>
<td>0.564</td>
</tr>
</tbody>
</table>

It can be seen that when the k value is 3, the GA index reaches the maximum. Therefore, for the data set, the optimal number of clusters is 3. At this time, the clustering model has stronger generalization ability and is the best aggregation for the data set.

3.3. Feature Extraction

Calculate all the characteristics of the student through a certain feature evaluation function to get the score of each feature. Sort the features according to the score. The top several features can be used as the feature words of the student’s quality. This process is called feature extract. Common methods for feature extraction include TF-IDF algorithm, word frequency method, mutual information, expected cross entropy, information gain method, and chi-square statistics. This paper adopts the TF-IDF algorithm for clustering, and the clustering framework is shown in Figure 3:
3.4. Statistics

All data analysis in this article uses SPSS19.0, statistical test uses two-sided test, significance is defined as 0.05 and \( p < 0.05 \) is considered significant. The statistical results are displayed as mean±standard deviation (x±SD). When the test data complies with the normal distribution, the double T test is used as the comparison within the group, and the independent sample T test is used as the comparison. If the regular distribution is not sufficient, two independent samples and two related samples will be used for inspection.

4. Experimental Analysis of College Students’ Sports Core Literacy Education

4.1. Student Understanding

We conduct statistics on the core literacy of physical education, physical cognition, physical awareness, and physical participation. Before the experiment, we briefly introduce the basic structure of these three data sets. The sports cognition data set has 343 samples, the sample dimension is 6, and the correct class number is 2; the sports awareness data set has 762 samples, the sample dimension is 8, the correct class number is 2, and the sports parameter data set consists of 702 Sample, the sample dimension is 9, and the correct number of classes is 2. For these three data sets, DBSCAN algorithm, CLIQUE algorithm and the GA index proposed in this paper are used to evaluate the effectiveness of the clustering results, and the best number of clusters evaluated by each index is used as the standard to measure the quality of each index. The results are shown in Table 2:

<table>
<thead>
<tr>
<th>Data set</th>
<th>Number of samples</th>
<th>Sample dimension</th>
<th>Correct number of classes</th>
<th>Optimal number of clusters (DBSCAN)</th>
<th>Optimal number of clusters (CLIQUE)</th>
<th>Optimal number of clusters (GA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sports cognition level</td>
<td>343</td>
<td>6</td>
<td>2</td>
<td>16</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Sports awareness</td>
<td>762</td>
<td>8</td>
<td>2</td>
<td>21</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Sports participation</td>
<td>702</td>
<td>10</td>
<td>2</td>
<td>17</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>
Through the comparison of the clustering effectiveness indicators in Table 2 and the experimental results table, it can be found that the GA indicator can accurately find the optimal number of clusters for these three data sets, while the traditional DBSCAN indicator cannot get an accurate number of clusters for each data set. The CLIQU index can only accurately find the cluster number of sports awareness data sets, but it cannot accurately find the best cluster number for sports awareness and sports parameter data sets.

The most basic literacy of sports literacy is sports-related knowledge, that is, the degree of awareness of sports literacy. The most basic literacy of sports literacy is sports-related knowledge, that is, the degree of awareness of sports literacy. Using the method of random selection, 10 related concepts are selected, which are "health", "physical activity", "physical fitness", "lifestyle", "aerobic exercise", "exercise deficiency", "extreme" and "extreme". "BMI", "sub-health" and "exercise prescription", these concepts are highly related to sports terminology. Therefore, examining college students' awareness of them through them can fully reflect their awareness of sports terminology. The result is shown in Figure 4:

![Figure 4: Cognition of sports terminology](image)

It can be seen from Figure 4 that the respondents have the highest degree of understanding of "health", accounting for 55.2%, followed by "sports activities", "physique", and "lifestyle" accounting for 46.2%, 45.6%, and 39.2%, respectively. People are more familiar with these concepts mainly because they appear more frequently in life. The three concepts least understood are "BMI", "sub-health", and "exercise prescription", accounting for 33.0%, 24.2% and 28.8% respectively. The reason why everyone does not understand is that these concepts are all professional terms. Deliberate memory is hard to understand.

A large part of the role of sports is to make the body healthier, and fitness plays a healthy role. The following 20 fitness-related judgment questions are selected. Each question is answered correctly and no score is scored. The full score is 40 points. In order to analyze the student's score more intuitively, the results of converting the student's score into a hundred-point system are shown in Table 3:
Table 3: College students' awareness of sports knowledge

<table>
<thead>
<tr>
<th></th>
<th>90-100</th>
<th>80-89</th>
<th>70-79</th>
<th>60-69</th>
<th>&lt;60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of people</td>
<td>43</td>
<td>332</td>
<td>435</td>
<td>217</td>
<td>70</td>
</tr>
<tr>
<td>Percentage</td>
<td>3.92</td>
<td>30.36</td>
<td>39.56</td>
<td>19.95</td>
<td>6.21</td>
</tr>
</tbody>
</table>

It can be seen from the above table that most students' scores are concentrated between 70-89. Everyone has a good level of knowledge about fitness, and the failure rate only accounts for 6.20%.

4.2. Student Differences

The differences between students will greatly affect the effect of clustering. Therefore, we have made statistics on the differences between students and teachers. The differences between students and teachers are shown in Table 4, and the differences between teachers are shown in Table 5.

Table 4: Student differences

<table>
<thead>
<tr>
<th>Evaluation degree</th>
<th>Number of people</th>
<th>Differences</th>
<th>percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>No foundation</td>
<td>47</td>
<td>2.93</td>
<td>7.8%</td>
</tr>
<tr>
<td>Learn something about it</td>
<td>329</td>
<td>3.15</td>
<td>54.8%</td>
</tr>
<tr>
<td>Basic Mastery</td>
<td>179</td>
<td>3.98</td>
<td>29.8%</td>
</tr>
<tr>
<td>Mastery</td>
<td>45</td>
<td>3.8</td>
<td>7.5%</td>
</tr>
<tr>
<td>Not good</td>
<td>98</td>
<td>4.59</td>
<td>16.3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Self-assessment of physical fitness</th>
<th>Number of people</th>
<th>Differences</th>
<th>percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>commonley</td>
<td>245</td>
<td>5.17</td>
<td>40.8%</td>
</tr>
<tr>
<td>very good</td>
<td>196</td>
<td>5.53</td>
<td>33%</td>
</tr>
<tr>
<td>unclear</td>
<td>61</td>
<td>5.86</td>
<td>10.1%</td>
</tr>
</tbody>
</table>

Table 5: Differences in teacher evaluation

<table>
<thead>
<tr>
<th>Differences in physical fitness</th>
<th>Number of people</th>
<th>percentage</th>
<th>Differences in motor skills</th>
<th>Number of people</th>
<th>percentage</th>
<th>Differences in sports interests</th>
<th>Number of people</th>
<th>percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>very different</td>
<td>8</td>
<td>10.6%</td>
<td>17</td>
<td>22.7%</td>
<td>9</td>
<td>12%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>big difference</td>
<td>19</td>
<td>25.3%</td>
<td>28</td>
<td>37.3%</td>
<td>12</td>
<td>16%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The difference is average</td>
<td>32</td>
<td>42.7%</td>
<td>14</td>
<td>18.7%</td>
<td>35</td>
<td>46.7%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the table, we can see. In the self-evaluation of students and the evaluation of teachers, the differences between students are very large. The vast majority of students do not fully grasp the core literacy of physical education, and only less than 10% of the students believe that they have mastered it. This is physical education. The cluster analysis of core literacy education provides the basis. Through the statistics of students in different grades, the results are shown in Figure 5:
It can be seen from the figure that the average difference in the understanding of sports terminology among students of each grade is not too great. The freshman survey average was 27.53, with a standard deviation of 7.03; the sophomore survey average was 26.93, with a standard deviation of 6.77; the junior survey average was 27.00, with a standard deviation of 6.66. In order to find out whether there are significant differences in the degree of understanding of sports terminology among junior high school students in each grade, a one-way analysis of variance was further used. The results are shown in Table 6:

<table>
<thead>
<tr>
<th></th>
<th>Sum of squares</th>
<th>Df</th>
<th>Mean square</th>
<th>T</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>79.65</td>
<td>2</td>
<td>39.82</td>
<td>0.851</td>
<td>0.427</td>
</tr>
<tr>
<td>Within group</td>
<td>51214</td>
<td>1094</td>
<td>46.81</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>total</td>
<td>51293</td>
<td>1096</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The results in the table show that the P value of junior high school students in each grade in terms of understanding of sports terminology is 0.427 greater than 0.05. Therefore, there is no significant difference in the degree of understanding of sports terminology among students of each grade.

4.3 Changes in Core Literacy

We teach the selected classes and compare their changes before and after the cluster teaching to understand the effect of the cluster teaching on the core literacy of sports. The specific data is shown in Figure 6:
It can be seen from Figure 6 that after the clustering algorithm, college students’ interest in sports is significantly higher than before classification. The overall increase in traditional teaching is about 4%, and the largest increase after clustering is 10.3%. The effect is very obvious. We also made statistics on teaching quality, as shown in Figure 7:

**Figure 7: Students' sports scores before and after the experiment**

It can be seen from Figure 7 that before the clustering, the students' sports core literacy teaching scores remained basically unchanged. After the adoption of clustering, the students' physical performance has been greatly improved, the excellent rate has increased by about 10%, and the failing rate has dropped to 0, indicating that cluster analysis can play a role in practice.
5. Conclusion

Schools are places for educating people and have important responsibilities in cultivating students' core physical literacy. However, in order to improve students' core physical literacy, the cooperation of the students themselves is needed, and the students need to cultivate this awareness spontaneously. This requires the guidance of the school. Students are very unfamiliar with this concept in the initial stage of cultivating core physical literacy. Teachers play a very important role. Teachers should educate them to cultivate the ability and awareness of core physical literacy. With this ability and awareness, you can spontaneously cultivate your own core literacy, so that teachers and students work together to play a key role in cultivating the core physical literacy of students. This article also has some shortcomings. Due to time and funding, the sample of the experiment is limited to some students in the city’s colleges and universities, which may result in students with some characteristics not being counted, and there are still imperfections in the implementation content. The experiment should be expanded. Scope, conduct further scientific and comprehensive research to increase its effectiveness. In the future, it can be studied whether it can be extended to the teaching of other subjects.

Acknowledgements

2020 Changzhou University Curriculum Ideological and Political Teaching Reform Project, Project No: KCSZ202012, "College Physical Education" Curriculum Core Literacy Reform Pilot Project

References

[8] Ari, Abdil|Çaglayan, Hakan Salim. Relationship between Organizational Justice Perceptions and Organizational Commitment Levels of School of Physical Education and Sports


