Research on Higher Education Based on Factor Analysis

Kun Deng(✉), Yanhui Xue, and Dihui Deng

Hohai University, No. 200, Jinling North Road, Changzhou, China
1220967843@qq.com

Abstract. Today's times are changing rapidly. To adapt to the trend of the times, society needs a variety of high-level talents. As a result, colleges and universities that train these talents have gradually become the object of attention. Countries around the world are very concerned about the health of their higher education systems. Therefore, today we are here to check the pulse and temperature of higher education. We selected 6 s-level indicators and 10 third-level indicators to construct the Factor Analysis Model, and applied our model to conduct a comprehensive evaluation of higher education in 19 countries. Based on the results, we chose Pakistan with the lowest overall score as our targeted country whose system of higher education has room for improvement.

Keywords: Factor Analysis · System of higher education · Comprehensive evaluation

1 Introduction

1.1 Background

As the sustainable development agenda has gradually become an international policy, education has become an important focus based on achieving change [1], and in today's era of rapid development and change, society's demand for highly educated talents has surged. According to statistics from the UNESCO Institute for Statistics, the global gross tertiary enrollment rate increased from 10% in 1972 to 32% in 2012, and is now increasing by 1% every year. By 2014, the gross enrollment rate of higher education in 64 countries reached 50%, while only 5 countries reached that rate 20 years ago; 14 countries have the gross enrollment rate of higher education in t-he world over 80%. It can be seen that the importance of higher education in various countries has generally increased. It also coincides with the special period of the epidemic. How countries can adjust to achieve a healthy and sustainable higher education system and promote the further development of higher education has become a topic of practical significance for scholars to study and explore.

Z. Zhan et al. (Eds.): ICBDIE 2022, AHCS 5, pp. 1249–1256, 2023.
https://doi.org/10.2991/978-94-6463-034-3_129

1.2 Our Work

We establish a comprehensive evaluation model for higher education. We selected 6 s-level indicators and 10 third-level indicators to construct the Factor Analysis model, and applied our model to conduct a comprehensive evaluation of higher education in 19 countries. Eventually we chose Pakistan as our targeted country whose system of higher education has room for improvement based on your analysis.

2 Preparation of the Models

2.1 Assumptions and Justifications

BY adequate analysis of the problem, to simplify our model, we make the following well-justified assumptions. Other assumptions may not be as follows, but will be put forward in the model push.

- The indicator we selected for the latest year can roughly reflect the overall situation of the indicator.
- The 10 indicators we have selected can fully reflect the health and sustainability of hig-her education.
- All the data sources of the relevant documents inquired are accurate, reliable, stable and scientific.
- The 19 countries mentioned in the article represent all the countries to be analyzed in the world.
- It is assumed that there is no interaction relationship among the indicators in the health evaluation index system of higher education.
- The analysis process is not affected by extreme data, and there will be no abnormal results in the overall evaluation of the entire country due to the lack of an indicator.
- It is assumed that the countries in the electoral district will not undergo unusually large changes in the next 10 years.

2.2 Notations

See Table 1.

Table 1. Notation

Symbol	Description
x_i	Evaluation index
F_i'	Score on each common factor
μ	Average value
F_c	Overall ratings

(*continued*)

Table 1. (*continued*)

Symbol	Description
μ_i	Variance contribution rate
d	Factor loading
ε	Special factor
$\lambda(k)$	Series ratio
E	Residual sequence

3 Comprehensive Evaluation of Higher Education System Model

3.1 Confirmation of Indicators

Many researchers have discussed the evaluation of higher education. For example, Alexander W. Astin (1991) believes that the index system for evaluating the development level of higher education includes opportunities to receive and participate in higher education (participation rate in high school education, enrollment rate in higher education, etc.) [2]. Others believe that the index system for evaluating the development level of higher education mainly includes the population of higher education, financial investment in higher education and human resources (the proportional relationship between higher education financial expenditure and gross national product, etc.) [3]. We believe that health focuses on the present and sustainability focuses on the future, so statistics are made in the current and future directions, totaling 10 indicators. In the future, we should pay attention to the opportunities of higher education. Here we select representative indicators as the enrollment rate of college students and the total number of colleges and universities; the level of education will affect the cultural quality of the labor force and accordingly affect the future development of higher education. Here we select the representative indicator as Mean Years of Schooling. At the same time, the quality of education also affects the future development of higher education. The quality of education directly affects the number of students attracted.

3.2 Factor Analysis Method

We assume that the GDP growth steadily and the mathematical model of factor analysis is expressed as:

$$X_i = \mu_i + a_{i1}F_1 + \cdots + a_{iq}F_q + \varepsilon_i \tag{1}$$

where, $X = (x_1, x_2, \cdots x_p)$ is a p-dimensional random vector, μ is the average value. $F = (f_1, f_2 \cdots f_q)$ is a p-dimensional random vector, called a common factor [4], which is a common factor that affects the entire X;ε is a special factor, which means the part that cannot be explained by the common factor. $E(F) = 0, E(\varepsilon) = 0; \mathrm{Cov}(F, \varepsilon) = 0; D(F) = I$.

Table 2. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.513
Bartlett's Test of Sphericity	Approx. Chi-Square	130.151
	df	45
	Sig.	0.000

We select 19 representative countries, and through the search of relevant data.When selecting countries, we fully considered factors such as latitude, regional characteristics, economic level, and cultural development to make it more comprehensive and representative.

3.3 Implementation and Results

First, judge whether the original variable to be analyzed is suitable for factor analysis.Since factor analysis requires a strong correlation between variables, we must first analyze and test the original variables before performing factor analysis. We used SPSS statistical software to calculate the collected data and got the KMO test value of 0.513. Meanwhile the Bartlett's sphere test significance probability of 0.000 as shown in Table 2. According to the principle that the KMO test value is greater than 0.5 and the significance probability value of the Bartlett sphere test is less than 0.05, we believe that there is a strong correlation between the indicators, which is suitable for factor analysis.

Second, extract the main factor. According to the principle that the eigenvalue is greater than 1, three main factors are extracted using the principal component method, and the cumulative variance contribution rate is 78.919%, representing most of the information, indicating that the three main factors can more fully reflect the development of higher education. The results of extracting the principal factors are shown in Table 3 by SPSS statistical software.

Third, perform factor rotation. The degree of correlation between the main factor and the original variable index is characterized by the factor loading value. The larger the factor loading value, the more information the factor contains in the corresponding original variable index [5].

Next, build a factor score model. We use SPSS statistical software to directly obtain the component score coefficient matrix of the three main factors. The factor score model is established based on the score coefficient of each main factor, and the scores on each common factor F1, F2 and F3 are F1', F2', F3', respectively. The factor score function is as follows:

$$\begin{cases} F_1' = -0.044x_1 - 0.092x_2 - 0.165x_3 + 0.088x_4 + 0.228x_5 \\ +0.302x_6 + 0.324x_7 + 0.199x_8 + 0.192x_9 - 0.067x_{10} \\ F_2' = 0.216x_1 + 0.423x_2 + 0.019x_3 + 0.122x_4 - 0.029x_5 \\ -0.113x_6 - 0.098x_7 + 0.021x_8 - 0.360x_9 + 0.397x_{10} \\ F_3' = 0.259x_1 - 0.093x_2 + 0.629x_3 + 0.146x_4 + 0.047x_5 \\ -0.049x_6 - 0.295x_7 - 0.040x_8 + 0.282x_9 + 0.003x_{10} \end{cases} \tag{2}$$

Table 3. Total Variance Explained

Component	Initial eigenvalue variance percentage (Percentage of components)			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	4.996	49.962	49.962	4.996	49.962	49.962	3.807
2	1.575	15.749	65.711	1.575	15.749	65.711	2.372
3	1.321	13.208	78.919	1.321	13.208	78.919	1.713
4	0.702	7.018	85.937				
5	0.549	5.487	91.423				
6	0.465	4.646	96.069				
7	0.157	1.575	97.644				
8	0.121	1.209	98.854				
9	0.088	0.885	99.738				
10	0.026	0.262	100.000				

Finally, establish Comprehensive Evaluation of Higher Education System Model.In order to express the contribution of each public factor to the development of higher education, we weighted it according to the variance contribution rate of each public factor to the cumulative variance contribution rate, and established Comprehensive Evaluation of Higher Education System Model on this basis:

$$Fc = \frac{F_1'\mu_1 + F_2'\mu_2 + F_3'\mu_3}{\mu_1 + \mu_2 + \mu_3} \tag{3}$$

where, F_c is the comprehensive score, c is the subscript of F,and μ_i is the variance contribution rate of the common factor.Substituting the values of $\mu 1$, $\mu 2$, and $\mu 3$ in Table 4 into the above formula, we get:

$$Fc = 0.633F_1' + 0.1996F_2' + 0.1674F_3' \tag{4}$$

Then substituting the F1', F2', and F3' corresponding to each country obtained before into the above formula, we can get the comprehensive score and ranking, as shown in Table 4. The higher the overall score, the better the health of the higher education system. A positive value indicates that its condition is good and above average; a negative value indicates that its condition is bad and below average.

Table 4. National Higher Education System Health Status Score and Ranking

Country	F_1'	Ranking	F_2'	Ranking	F_3'	Ranking	F	Ranking
United States	0.66	5	2.42	1	-0.26	11	0.86	1
Sweden	0.82	3	-0.33	11	1.63	1	0.73	2
Germany	0.71	4	0.31	5	0.98	5	0.67	3
Japan	0.55	6	1.31	3	-0.02	9	0.61	4
Israel	1.3	1	-0.98	16	-0.72	13	0.51	5
United Kingdom	0.25	8	0.25	7	1.36	3	0.44	6
France	0.51	7	-0.06	9	0.7	6	0.43	7
Australia	0.92	2	-0.55	13	-0.58	12	0.37	8
Malaysia	-0.12	11	-0.41	12	0.53	7	-0.07	9
China	-0.54	15	1.58	2	-0.77	14	-0.16	10
Russia	-0.09	10	0.3	6	-1.12	17	-0.19	11
South Africa	-0.34	13	-1.14	18	1.22	4	-0.24	12
Brazil	-0.28	12	-0.76	15	0.35	8	-0.27	13
Argentina	0.15	9	-1.27	19	-0.93	15	-0.31	14
Vietnam	-0.5	14	-1.04	17	-0.16	10	-0.55	15
India	-1.63	19	1.05	4	1.41	2	-0.59	16
Thailand	-0.55	16	-0.7	14	-0.94	16	-0.64	17
Iran	-0.76	17	0.12	8	-1.52	19	-0.71	18
Pakistan	-1.05	18	-0.1	10	-1.19	18	-0.88	19

4 Conclusions

Based on the Comprehensive Evaluation of Higher Education System Model, combined with Table 4 and the original data, the health status of the higher education system in various countries can be evaluated.

In the first main factor, the top 3 scores are Israel, Australia, and Sweden. According to the data,they are all developed countries with super-high humanities, with high incomes and large enrollment opportunities, so the labor force has a high level of cultural education. The three countries with the lowest scores are Iran, Pakistan, and India, which are mainly distributed in the southwestern corner of Asia. The level of higher education development is relatively backward, and the penetration rate and admission opportunities are far below average.

In the second main factor, the top 3 scores are the United States, China, and Japan. According to data from the World Bank, they are all powerful economies in the world, with a large number of colleges and universities and a large scale of higher education. The three countries with the lowest scores are Vietnam, South Africa, and Argentina, which have relatively limited higher education resources and weak foundations.

In the third main factor, the top 3 scores are Sweden, India, and the United Kingdom. These countries have an absolute advantage in terms of TSS. The three countries with the lowest scores are Russia, Pakistan, and Iran. These countries have relatively low investment in higher education, which has affected the health and sustainability of the higher education system in various countries to a certain extent.

Among the comprehensive factors, eight countries have a comprehensive score greater than zero. The top 3 countries with a score are the United States (score 0.86), Sweden (score 0.73), and Germany (score 0.67). They are all developed countries, super-high humanities development countries, high economic levels, and rich higher education resources. Among them, the United States has the highest score in the second main factor (score 2.42). The scale of development of higher education is relatively large. The overall scores of the remaining 11 countries are all below zero, which indicates that the health status of the higher education system in these countries is below average, and it also shows that the development of higher education in the world is not balanced. Further analysis found that the lowest overall score of Pakistan was -0.88, and even the scores of the first and third factors were lower than -1, ranking in the top 2 of the lowest scores. This shows that the health status of Pakistan's higher education system is in a poor state.

In summary, we select Pakistan as a country where the higher education system can be improved.

References

1. Liu Mengyue. Research on the Status Quo of Sustainable Development of Asian Higher Education[J]. Examination Weekly, 2013(20):17.
2. Alexander W. Astin. Assessment for Excellence :e Philosophy and Practice of Assessment and Evaluation in Higher Education [M]. Westport: the Oryx Press, 1991 :38 -94.
3. Chu Jiangting, Guo Dexia. A Probe into the Development Index System of Higher Education in my country [J]. Jiangsu Higher Education,2003(1) :24 -25.
4. Wang Xiaojing.Factor analysis and comprehensive evaluation of the development level of high-tech industries in my country's provinces (cities, autonomous regions) [J]. Mathematical Practice and Understanding,2007(18) :19.
5. Zhu Yongdong, Xiang Xinghua, Ye Yujia.Research on Comprehensive Evaluation of American Higher Education Development Level Based on Factor Analysis [J]. Higher Education Exploration, 2014(5):70