Research on Teaching Quality Based On Principal Component Analysis

Jing Han[1], Jing Ma[2]

1*School of Foreign Languages*
Handan College
Handan, Hebei, China
2School of P. E.,
Handan College,
Handan, Hebei, China

***Abstract* — in this paper we use principal component analysis to get factors affecting the quality of English teaching, and then construct comprehensive evaluation system. On this basis, we use fuzzy theory to evaluate the data by three-level indicator. Finally we use normalization method to get the result, which provides the scientific theory for English teaching quality and improves the quality of teaching in English.**

Keywords - Principal Component Analysis; Teaching Quality; Fuzzy Comprehensive Evaluation

I. INTRODUCTION

English teaching is an important part of Chinese universities. But now there are problems in English teaching, for example, English teaching time is long but the effect is poor, lots of students have good exam results but they have low ability to use English and so on. So, college English teaching work must conform to the demands of society and the development of the times, and improve English teaching effectiveness and efficiency.

In order to improve the quality of English teaching, strengthen teaching management and further mobilize the enthusiasm of English teachers, we need quantitative method to evaluate English teaching. In order to fully assess teaching, this paper chooses appropriate English teaching indicators. Because lots of indicator may cause complex operation and information overlapping. So this paper chooses a small amount of non-relevant indicators to replace lots of related indicators. At the same time the small amount non-relevant indicators can reflect the original index information. In the current, evaluation of teaching ability of English teacher is ultimate assessment (only use students' score to evaluate teacher). How to analyze students' final grades can objectively reflect the teacher's teaching? This paper considers absolute indicator and relative indicator which can reflect English proficiency of students. Next we use principal component analysis to analyze these selected indicators and then get comprehensive evaluation value of teacher education. In addition to considering the student's final grade, it will also take into account other data. So, by questionnaire the paper gets some indicators divided into main and sub factors. Because these indicators contain ambiguity, so the quantitative analysis is not easy. Then the paper constructs multi-level fuzzy comprehensive evaluation model.

II. THE COMPREHENSIVE EVALUATION INDEX SYSTEM

A. Principle Introduction

Principal component analysis as a method of statistical analysis converts original indicators into several comprehensive indicators. Its main function is to reduce dimension. Assume there are n geographic samples, each sample has p variable descriptions. Then we get $n \times p$ geographic data matrix:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ M & M & \cdots & M \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

How to find out the internal regularity of geographic thing from lots of variable data? The answer is that we solve this problem in p -dimensional space, but this kind of practice is very trouble. So we must reduce the dimension, that is using a few comprehensive indicators to replace original many indicators. However how to obtain these comprehensive indicators? The simplest solution is to use linear combination of the original indicators.

Let $x_1, x_2, \cdots, x_i$ be the original variable indicators, $x_1, x_2 \cdots z_m (m \le p)$ are their comprehensive indicators (i.e. new variable indicators), then

$$\begin{cases} x_1 - l_{11}x_1 + l_{12}x_2 + \cdots + l_{1p}x_p \\ x_2 - l_{21}x_1 + l_{22}x_2 + \cdots + l_{2p}x_p \\ \cdots\cdots \\ z_m - l_{m1}x_1 + l_{m2}x_2 + \cdots + l_{mp}x_p \end{cases}$$

z_1 is linear combination of $x_1, x_2, \cdots, x_p$ satisfying biggest variance; z_2 is linear combination of $x_1, x_2, \cdots, x_p$

satisfying biggest variance and unrelated with z_1 ; z_m is linear combination of $x_1, x_2, \cdots, x_p$ satisfying biggest variance and unrelated with $z_1, z_2, \cdots z_{m-1}$.

The new indicator $z_i\ (i = 1,2,\cdots,m)$ is the i-principal component of original indicator $x_j\ (j = 1,2,\cdots,p)$. The largest percentage in total variance is z_1 , and then variance of $z_2, z_3, \cdots, z_m$ are successive reduction. We often choose a few in the front of the maximum principal components. The aim is to reduce the index, but also made clear the main contradiction, simplify the relationship between indicators.

Based on the above analysis, to obtain the principal component is that we determine the load $l_{ij}(i = 1,2,\cdots,m;\ j = 1,2,\cdots,p)$ of original indicator $x_j(j = 1,2,\cdots,p)$ in the principal component $z_i(i = 1,2,\cdots,m)$. They are eigenvectors corresponding to m larger eigenvalues of correlation matrix of $x_1, x_2, \cdots, x_p$ respectively.

B. System Construction

We put all individual data in the original data into the above every linear combination formula of the principal component, and then get principal component score values of all main bodies. The absolute indicator related with students' n times English exam scores are the following:

Average score reflecting the overall level of teachers' class:

$$\overline{c} = \frac{\text{total grade of n times exam}}{\text{total number of participating n times exam}}$$

Variance reflecting the separation degree of teachers' class:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(c_i - \overline{c})^2}$$

Pass rate:

$$\alpha = \frac{\text{number of passing the n times exam}}{\text{number of participating the n times exam}}$$

Excellent rate:

$$\beta = \frac{\text{excellent people number of n times exam}}{\text{excellent number of participating n times exam}}$$

The relative indicator is the progress rate

$$\gamma = \sum_{i=1}^{n-1}\frac{E_i}{N_i}.$$

Through the principal component analysis to analyze the five indicators, we can obtain objectively English teacher's teaching comprehensive evaluation indicator E . In this paper, we randomly select four classes A, B, C, D from any university. The four classes' principal component factors are the following:

progress rate

$x_1 = (0.666667, 0.454545455, 0.675676, 0.774194)$

pass rate

$x_2 = (0.930555556, 0.909090909, 0.932432432, 0.983870968)$

excellent rate

$x_3 = (0.152777778, 0.227272727, 0.135510511, 0.24935484)$

average score

$x_4 = (74.59722222, 76.25, 73.32432432, 76.67741935)$

standard deviation

$x_5 = (5.103538321, 8.8705975, 6.480966119, 5.139267531)$

First, we calculate x_{ij} , and then get the original data matrix X :

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}^T = \begin{bmatrix} 0.66667 & 0.93056 & 0.15278 & 74.597 & 5.1035 \\ 0.45455 & 0.90909 & 0.22727 & 76.25 & 8.8706 \\ 0.67568 & 0.93243 & 0.13551 & 73.324 & 6.481 \\ 0.77419 & 0.98387 & 0.24194 & 76.677 & 5.1393 \end{bmatrix}$$

Second, we remove the dimensional interference between indicators. By the standardized treatment, we get standardized matrix Z where $Z_{ij} = \frac{x_{ij} - \overline{x_i}}{s_j}$. Further we construct the characteristic equation of relative matrix v of x_i and x_j , then get the characteristic root $(\lambda_1, \lambda_2, \cdots \lambda_p)$. Through $q_i = \frac{\lambda_i}{\sum \lambda_i}$, we get the contribution rate q_i .

TABLE I COMPOSITION ANALYSIS

Composition analysis	characteristic root λ_i	contribution rate q_i
$(\chi_1)1$	2.36259	0.47252
$(\chi_2)2$	1.35642	0.27128
$(\chi_3)3$	1.281	0.2562
$(\chi_4)4$	0	0
$(\chi_5)5$	0	0

Next, by Jacobi iteration method of orthogonal transformation, we get the corresponding eigenvectors matrix

$$\begin{bmatrix} 0.29714 & 0.47187 & -0.43607 & -0.64782 & 0.28146 \\ -0.41695 & 0.63211 & 0.065738 & -0.056332 & -0.64738 \\ 0.73834 & -0.16843 & -0.083372 & -0.0092981 & -0.64765 \\ 0.25441 & 0.47054 & -0.31631 & 0.75816 & 0.1975 \\ 0.35776 & 0.35776 & 0.8\ 3577 & -0.047567 & 0.20791 \end{bmatrix}$$

Finally, by Matlab we get the score matrix

$$\begin{bmatrix} -1.6257 & -0.065298 & -0.019943 & -1.711 \\ 0.11259 & 0.78723 & -0.33308 & 0.56674 \\ 0.17744 & 0.0092437 & 0.85307 & 1.0398 \\ 0.40092 & -0.73449 & -0.42562 & -0.75919 \end{bmatrix}$$

We can calculate E(A)= -3.4219, E(B)= 1.1335, E(C)= 2.0796 and E(D)= -1.5184 from score matrix. So A class is the worst, next is D class, next is B class and C class is the best.

III. MULTI-LEVEL FUZZY COMPREHENSIVE EVALUATION MODEL

A. *Theory Introduction*

Assume k layer compose factor domain U（$k \geq 2$）, $U = (U_1^{(1)} \quad U_2^{(1)} \quad \cdots \quad U_m^{(1)})$ is m-th factor of the first layer (the highest layer), $V = (v_1 \quad v_2 \quad \cdots \quad v_n)$ is its review set, so the multi-level fuzzy comprehensive evaluation model is $B = A \circ R$ where A is each layer's weight vector, x is every weight vector of the x+1 layer, R is the lowest layer (the k layer) fuzzy relation matrix.

Multi-level fuzzy comprehensive evaluation method is that we calculate from the lowest layer (the k layer) upward step by step until to get the last set of review B. The evaluation conclusion of the k layer is membership of the $k-1$ layer factor. Calculation steps are

(1) calculate the fourth layer and get

$$B_{111} = A_{111} \circ R_{111}$$
$$B_{112} = A_{112} \circ R_{112}$$
$$\vdots$$
$$B_{11s} = A_{11s} \circ R_{11s}$$
$$\vdots$$
$$\vdots$$
$$B_{mq1} = A_{mq1} \circ R_{mq1}$$
$$B_{mq2} = A_{mq2} \circ R_{mq2}$$
$$\vdots$$
$$B_{mqo} = A_{mqo} \circ R_{mqo}$$

When finishing the calculation of the third layer, let

$$R_1 = \begin{pmatrix} B_{11} \\ B_{12} \\ \vdots \\ B_{1p} \end{pmatrix}, \cdots, R_m = \begin{pmatrix} B_{m1} \\ B_{m2} \\ \vdots \\ B_{mp} \end{pmatrix}.$$

(2) calculate the second layer and get

$$B_1 = A_1 \circ R_1$$
$$\vdots \quad .$$
$$B_m = A_m \circ R_m$$

When finishing the calculation of the second layer, let

$$R = \begin{pmatrix} B_1 \\ \vdots \\ B_m \end{pmatrix}.$$

(3) calculate the highest layer and get the last set of review $B = A \circ R$, then perform quantitative processing.

B. *Model Establishment*

Now commonly used way of evaluation is the ultimate evaluation, that is using score to evaluate English teacher. So under this method how to analyze the student's final grade can objectively reflect the teacher's teaching? In this paper, we consider absolute indicator and relative indicator which can reflect students' English level. Next we use principal component analysis to analyze these selected indicators and then get comprehensive evaluation value of teacher education. In addition to considering the student's final grade, it will also take into account other data. So, by questionnaire the paper gets some indicators divided into main and sub factors. Because these indicators contain ambiguity, so the quantitative analysis is not easy. Then the paper constructs multi-level fuzzy comprehensive evaluation model.

First, build on questionnaire about English teachers' teaching quality and level, and then investigate the access, finally get the following data:

TABLE II ENGLISH EDUCATION TEACHING QUALITY EVALUATION INDEX SYSTEM

First layer	Second layer (main factor)	Third layer (sub factor)
Teaching quality	Teaching attitude F1（0.2）	The degree of class preparation F11（0.3）
		Classroom teaching record F12（0.3）
		Homework and after-school tutoring F13（0.4）
	Teaching content F2（0.3）	Complete the requirements of teaching outline F21（0.2）
		The combination of extracurricular materials and materials F22（0.4）
		The implementation of the class discussion F23（0.4）
	Teaching strategies and methods F3（0.2）	Arouse the students' interest in learning English F31（0.4）
		The moderation of English teaching F32（0.3）
		Guidance for differences F33（0.3）
	Teaching effect F4（0.3）	Test scores at ordinary times F41（0.4）
		English communication and application ability F42（0.4）
		Class attendance and homework F43（0.2）

TABLE III THE NUMBER OF STUDENTS AND CLASS TEACHER

Main factor	Sub factor	Student evaluation（200）					class teacher evaluation（4）				
		better	good	general	poor	poorer	better	good	general	poor	poorer
Teaching attitude F1（0.2）	F11(0.3)	46	24	26	4	0	1	1	0	0	0
	F12(0.3)	42	20	32	4	2	1	0	1	0	0
	F13(0.4)	20	18	50	10	2	0	0	1	1	0
Teaching content F2（0.3）	F21(0.2)	34	26	18	22	0	0	1	1	0	0
	F22(0.4)	18	29	36	17	0	0	0	1	1	0
	F23(0.4)	17	18	36	28	1	0	0	1	1	0
Teaching strategies and methods F3（0.2）	F31(0.4)	48	24	18	10	0	0	2	0	0	0
	F32(0.3)	40	28	30	2	0	1	0	1	0	0
	F33(0.3)	28	34	29	8	1	0	2	0	0	0
Teaching effect F4（0.3）	F41(0.4)	50	26	20	3	1	2	0	0	0	0
	F42(0.4)	16	24	31	25	4	0	0	2	0	0
	F43(0.2)	46	26	22	6	0	1	0	1	0	0

Second, analyze collected objects and fuzzy evaluate the third layer indicator.

TABLE IV FUZZY EVALUATION OF INDICATOR F11 FROM 200 STUDENTS

rank	better	good	general	poor	poorer
number	46	24	26	4	0
proportion	0.46	0.24	0.26	0.04	0.00

The evaluation result is fuzzy set R111=(0.46,0.24,0.26,0.04,0.00)。Similarly, we can get the following fuzzy set of indicator F12、F13 from 200 students

R112=(0.42,0.20,0.32,0.04,0.02)

R113=(0.20,0.18,0.50,0.10,0.02)

So we obtain the single factor evaluation matrix of indicator F1

$$R_{11}=\begin{bmatrix}0.46 & 0.24 & 0.26 & 0.04 & 0.00\\ 0.42 & 0.20 & 0.32 & 0.04 & 0.02\\ 0.20 & 0.18 & 0.50 & 0.10 & 0.02\end{bmatrix}$$

Next analyze collected objects and fuzzy evaluate the second layer indicator. The weight distributions of three indicators of teaching attitude F1 are F11(0.3)、F12(0.3)、F13(0.4), so A11=(0.3,0.3,0.4) is a fuzzy set. Finally the comprehensive evaluation of F1 from 200 students is

$$B11'=(0.3,0.3,0.4)\circ\begin{bmatrix}0.46 & 0.24 & 0.26 & 0.04 & 0.00\\ 0.42 & 0.20 & 0.32 & 0.04 & 0.02\\ 0.20 & 0.18 & 0.50 & 0.10 & 0.02\end{bmatrix}$$

=(0.3440 0.2040 0.3740 0.0640 0.0140)

We use "normalization" to process evaluation results B11'.

From 0.3440+0.2040+0.3740+0.0640+ 0.0140=1, we get

$$B_{11}=\left(\frac{0.3440}{1},\frac{0.2040}{1},\frac{0.3740}{1},\frac{0.0640}{1},\frac{0.0140}{1}\right)$$

=(0.3440 0.2040 0.3740 0.0640 0.0140)

From the normalization result, 34.40 percent of 200 students think that the evaluation of this teacher's teaching attitude is "better", 20.40 percent is "good", 37.40 percent is "general", 6.40 percent is "poor", 1.40 percent is "poorer". Similarly, we can get the comprehensive evaluation of F2, F3, F4

B12=(0.2008 0.24 0.324 0.24 0.004)

B13=(0.3960 0.2820 0.2490 0.0700 0.0030)

B14=(0.3560 0.2520 0.2480 0.1240 0.0200)

Then we get

$$R_1 = \begin{bmatrix} 0.3340 & 0.2040 & 0.3740 & 0.0640 & 0.0140 \\ 0.208 & 0.24 & 0.324 & 0.224 & 0.004 \\ 0.3960 & 0.2820 & 0.2490 & 0.0700 & 0.0030 \\ 0.3560 & 0.2520 & 0.2480 & 0.1240 & 0.0200 \end{bmatrix}$$

So

$$B1=A1 \circ R1=(0.2\ 0.3\ 0.2\ 0.3) \circ \begin{bmatrix} 0.3340 & 0.2040 & 0.3740 & 0.0640 & 0.0140 \\ 0.208 & 0.24 & 0.324 & 0.224 & 0.004 \\ 0.3960 & 0.2820 & 0.2490 & 0.0700 & 0.0030 \\ 0.3560 & 0.2520 & 0.2480 & 0.1240 & 0.0200 \end{bmatrix}$$

=(0.3172 0.2448 0.2962 0.1312 0.0106)

B1 is the teacher's fuzzy comprehensive evaluation. It shows that 31.72 percent of 200 students think that the evaluation of teacher is "better", 24.48 percent is "good", 29.28 percent is "general", 13.12 percent is "poor", 1.06 percent is "poorer". Similarly we get the fuzzy comprehensive evaluation result of class teacher

B2=(0.2400 0.2000 0.4000 0.1600 0.0000)

$$R_1 = \begin{bmatrix} 0.3150 & 0.2423 & 0.2928 & 0.1289 & 0.0210 \\ 0.2400 & 0.2000 & 0.4000 & 0.1600 & 0.0000 \end{bmatrix}$$

So $R = \begin{pmatrix} 0.4 & 0.6 \end{pmatrix} \circ$

$R_1^{'}$=(0.2700 0.2169 0.3571 0.1476 0.0084)

Now we give score to the rank: better——90~100; good——80~89; general——70~79; poor

——60~69; poorer——50~59。So we get the fuzzy comprehensive evaluation score of the English teacher

$$G = 0.27088 \times 95 + 0.21792 \times 85 + 0.35848 \times 75 + 0.14848 \times 65 + 0.00424 \times 55 = 81.027$$

It belongs to rank "good".

IV. CONCLUSION

In order to construct comprehensive evaluation indicator system for English teaching quality, the paper first obtain the five indicators: average score, variance, pass rate, excellent rate and progress rate, and then use principal component analysis to get the objective teaching quality comprehensive evaluation indicator. Next we test the model and finally get the comprehensive evaluation value.

In order to construct multi-level fuzzy comprehensive evaluation model for English teaching quality, the paper first get the evaluation indicator number of students and class teachers, and then get the teacher's evaluation rank by multi-level fuzzy comprehensive evaluation. The aim is promoting English teaching and improving teacher's teaching ability.

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, No. 2, pp. 529–551, 1955.

[2] Zhang Ningning, He Dengfa, Sun Yanpeng, et al., "The global giant carbonate field distribution characteristics and controlling factors of", Chinese petroleum exploration, vol. 2, No. 07, pp. 54-65, 2014.

[3] Qiu Liang, Haiwu Huang, "Principal component analysis in the empirical research of teaching quality evaluation", Education Teaching Forum, No. 41, pp.115-117, 2014.

[4] Li Zhang, Jiaoshi Yang, Hui Yuan, Lianggang Wang, "Research on Evaluation Methods of Teaching Quality of College Teachers", The Science Education Article Collects, No.12, pp.2-4, 2013.

[5] Lei Chen, Yan Wang, "The teaching quality evaluation model based on multivariate statistical analysis", Electronics World, No. 12, pp.199-199, 2013.

[6] Xiujuan Jiang, "The Fuzzy Optimization Model for College English Teaching Evaluation. Heilongjiang Researches on Higher Education", No. 11, pp.180-182, 2007.

[7] Qiuju Wang, Shicheng Zhang, "Application of Fuzzy Multifactorial Evaluation on Quality of College English Class Teaching", Journal of LANG FANG TEACHERS, vol. 8, No. 4, pp.38-39, 2008.