

Reliability Analysis of Students' Evaluation System -A Practical Study Based on the Evaluation Results of College Students

Li Lin^{1,a}, Wenping Deng^{1,b,*} and Rencai Zhou^{1,c}

¹Hubei University of Chinese Medicine, Wuhan, China

^a2832939734@qq.com, ^b1050513548@qq.com, ^c22320232@qq.com

*correspondence author

ABSTRACT

Objective is to study the consistency, stability and effectiveness of the evaluation index system of college students, and to analyze the reliability of the teaching quality evaluation system, and to provide a theoretical basis for improving the mechanism of "evaluating education by evaluation". The method expounds the theoretical basis of reliability analysis. Based on SPSS statistical analysis software, through the re-measurement reliability analysis method and the cloned Bach coefficient (α coefficient), the statistical analysis of the evaluation results of college students is carried out, and the reliability coefficient of the teaching quality evaluation system is calculated. The results obtained the correlation coefficient and the clone Bach coefficient of the evaluation scale. It is known from the calculation results that the evaluation scale has good correlation and good internal consistency. In Conclusion, the teaching quality evaluation system of the school is relatively stable and passed the internal consistency test, but it still needs to further improve the evaluation system of college students. The evaluation results can be maximized in the teaching of teachers, thus improving the quality of teaching.

Keywords: student evaluation, reliability, evaluation system, SPSS analysis

1. Introduction

With the development of contemporary society and the advancement of science and technology, the requirements for higher education have gradually increased. In recent years, China's major colleges and universities have begun to carry out large-scale reform of the teaching system, innovative teaching mode, and comprehensively improve the ability and level of talent training in colleges and universities. On how to promote the progress and development of higher education, the State's «Undergraduate Teaching Quality and Teaching Reform Project» states that improving teaching quality is the core issue of higher education, and teaching evaluation is a powerful means to improve teaching quality.^[1] At present, colleges and universities have established a variety of teaching evaluation methods, mainly peer evaluation, special expert evaluation and student evaluation. Among them, students are the most direct sensation to teachers' teaching and classroom activities, and participate in teaching activities throughout the whole process, which plays an important role in teaching evaluation. "Student Evaluation" is a kind of teaching quality management method that evaluates and evaluates the quality of teachers' classrooms by taking students as the evaluation subject, teachers and courses as evaluation objects. The wide implementation of the student evaluation system can actively promote the development of universities and the development of teaching, and has important practical significance for students to meet the learning needs to the maximum extent. With the in-depth study of the students' evaluation system and the importance attached to the evaluation results, major universities have applied the results of student evaluation to teachers' assessment and job evaluation, job promotion conditions, and the effectiveness of teaching

activities. In order to ensure the "evaluation by teaching" mechanism to achieve the expected results, the teaching evaluation results have practical reference value, to meet the needs of students to express learning, and to provide a basis for teachers to improve teaching results, to construct a set of sound and reasonable, in line with the professional characteristics of the school. It is extremely urgent to evaluate the teaching quality evaluation system and analyze and improve the reliability and validity of the teaching quality evaluation system.

2. Materials and Methods

2.1. Source

The research data of this paper is derived from the real evaluation data of teachers in the 2015-2018 school year of a university's information engineering college. Based on the teaching quality evaluation index system in the school's educational administration system, the data is collected from the teaching attitude, teaching content and teaching methods at the end of the semester. And the teaching effect is to grade the classroom quality of the teachers. In order to exclude other interference factors, the evaluation scores selected by the reliability analysis are the same for the same teacher, the same teaching method, the same teaching material, the same stage and the same evaluation mechanism. Evaluation data in the environment.

2.2. Statistical Analysis Methods

Reliability is reliability. Reliability analysis is the measurement of the consistency or stability of a measurement^[2] when measuring the same object through a measurement tool. A complete and sound teaching quality evaluation system, regardless of the measurement tools and measurement methods used in the research, the measurement results have a good consistency. The reliability analysis methods mainly include: retest reliability analysis method, replica reliability analysis method, half-fold reliability analysis method, and α reliability coefficient analysis method. The reliability analysis metric is usually expressed by the correlation coefficient. The larger the absolute value of the correlation coefficient, the higher the stability of the measured object. This paper mainly uses the test-retest reliability analysis method and the α -coefficient analysis method, and uses SPSS statistical analysis software to analyze the reliability of the student evaluation index system.

The retest reliability is the retest reliability, which is measured twice for the same measurement object using the same measurement method, and then the correlation coefficient^[3] of the two measurement results is calculated. The correlation coefficient is calculated as

$$r = \frac{N \sum XY - \sum X \sum Y}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

r represents the correlation coefficient of the two measurement results, X represents the score of the first measurement result, Y represents the score of the second measurement result, and N represents the number of people participating in the test.

The clone Bach coefficient is mainly used to test the internal consistency of the evaluation system and is suitable for

multi-score measurement data. It is the most commonly used information detection tool^[4] in current sociological statistics. The higher the alpha coefficient, the higher the reliability of the measured data. The formula for calculating the alpha reliability coefficient is:

$$\alpha = k [1 - (\sum_1^k S_i^2) / S_x^2] / (k-1)$$

Where k represents the number of tests, S_i Indicates the standard deviation of the i-th test score, S_x To test the standard deviation of the total score.

3. Results

3.1. Test-retest Reliability Analysis

The selected subjects selected 10 teachers from a college of information engineering, and obtained the results of the evaluation of the students in the second semester of 2016-2017 through certain technical means. The data was collected through the same teaching quality evaluation form at appropriate intervals. Conduct evaluations and statistics on the evaluation results. All the students tested have attended the classes of these 10 teachers. The statistics of the two evaluation results are shown in Table 1, where A, B, C, D, E, F, G, H, I, J represent 10 teacher.

Table 1 Statistics on the results of evaluation of 10 teachers

teacher	the first time	the second time
A	92.899	92.731
B	92.537	92.261
C	92.618	92.403
D	92.88	91.833
E	92.996	93.284
F	92.141	89.74
G	92.621	92.988
H	93.076	91.746
I	93.023	92.759
J	93.431	92.996

The results of the two-study teacher's two-semester evaluation were entered into the SPSS statistical analysis software, and the Pearson product-related correlation calculation method was used. The significance level was $\alpha=0.05$, and the analysis results are shown in Table 2.

Table 2 Pearson product difference calculation results

	the first time	the second time
Pearson Relative	1	0.655
Significance		0.040
N	10	10
Pearson Relative	0.655	1
Significance	0.040	
N	10	10

3.2. Clone Bach Coefficient Reliability Analysis

This paper evaluates the teacher's teaching effect and teaching methods through the evaluation index composed of the school's teaching quality evaluation system. The evaluation scores of 10

teachers in the 6-semester of 2015-2018 are selected as the objects to be measured, and the cloned Bach coefficient is calculated. The data is entered in the spss statistical analysis software. The statistical data of the measurement data is shown in Table 3, where A, B, C, D, E, and F represent 6 semesters between 2015 and 2018, respectively.

Table 3 α coefficient reliability detection data

A	B	C	D	E	F
90.96	93.08	93.28	93.40	92.97	92.99
92.42	94.31	92.76	92.94	92.94	92.87
90.87	94.11	92.99	92.85	92.62	92.84
92.37	93.67	92.00	92.31	92.74	92.51
91.20	93.20	93.97	93.26	93.43	93.04
92.30	92.48	92.49	92.85	92.90	92.86
91.17	91.36	89.74	92.17	92.13	92.50
91.14	91.85	91.75	92.74	92.33	92.48
93.01	91.81	92.37	92.86	92.28	92.62
90.50	91.68	91.68	92.71	93.00	93.26

The SPSS statistical analysis software is used to analyze the reliability of the measured data, and the α reliability coefficient of the measured data is detected, and the following analysis results are obtained. Table 4 is a statistical synthesis table between variables showing the average number of items and the average correlation coefficient between items.

Table 4 Summary Statistics

	Average	Maximum	Minimum	Range	Maximum/ Minimum	Variance	Number of Items
Item Average Number	92.490	91.594	92.797	1.203	1.103	0.226	6
Item Variance Number	0.592	0.069	1.312	1.312	18.917	0.298	6
Relevance Between Items	0.307	-0.426	0.789	0.789	-1.852	0.149	6

Table 5 is a graph showing the overall change trend after deleting a certain part. By calculating, the overall mean value, variance and other statistics of the remaining items when deleting certain types of data.

Table 5 State diagram after deleting a measurement data

	Average wide(if the item was deleted)	Variance wide(if the item was deleted)	Alternate after revision	Square Alternate	Cronbach Aplha(if the item was deleted)
A	463.3440	7.092	-0.015	0.628	0.750
B	462.1830	4.158	0.560	0.856	0.526
C	462.6350	3.065	0.839	0.963	0.337
D	462.1830	6.732	0.424	0.926	0.619
E	462.2040	6.293	0.646	0.811	0.576
F	462.1410	7.181	0.348	0.738	0.644

Tukey's additivity is used to test whether there is a significant interaction between the measured data of each column, and the variance analysis table of the project is obtained, indicating the difference between each group of the project, within the group, between the measurement data of each column, and the residual Items and non-additives, etc., and display F values and sig. The analysis of variance table is shown in Table 6.

Table 6 Analysis of variance table

	Sum of square	df	Mean Square	F	significance
Between staff	11.613	9	1.290		
Between projects	11.320	5	2.264	5.006	0.001
residual	0.010	1	0.010	0.021	0.886
Balance	20.343	44	0.462		
Sum	43.286	59	0.734		

The calculation results of the Reliability Statistics homogeneity reliability of the teaching quality evaluation scale can indicate the size of the Alpha value, and the confidence interval is 95%. The correlation coefficient table in the group is shown in Table 7.

Table 7 Group correlation coefficient table

	Intra-class correlation	95% Confidence Interval		F value test with True value 0			
		Lower Bound	Upper Bound	Value	df1	df2	significance
Single measurement	0.236	0.030	0.599	2.853	9	45	0.010
Average measurement	0.649	0.154	0.899	2.853	9	45	0.010

4. Conclusion

4.1. Test Reliability Analysis Results

Through the test of retest reliability test, it can be found that after the two evaluation results are calculated as Pearson

product difference, the correlation coefficient is 0.655. The experimental value is obtained at the level of significance of 0.05. It can be seen that * indicates that "the correlation between the two results is significant (bilateral)", that is, the reliability of the evaluation index system of the university students is 0.655 ($\alpha = 0.05$), and the correlation shows that the reliability is more reasonable and can be used as a classroom evaluation for college students. One of the means of quality and teacher teaching effectiveness.

4.2. Cloning Bach Coefficient Reliability Analysis Results

In this paper, the reliability analysis of the teaching quality evaluation table is carried out by cloning the Bach coefficient (α coefficient), and the correlation matrix, comprehensive statistical data, variance analysis table and correlation coefficient in the group are obtained. It can be seen from the correlation matrix that the part with the highest correlation coefficient is concentrated between D and F, and the correlation is between E and F, which is 0.789, which is highly correlated, which indicates that the correlation between the latter measurement data is more To be significant. In the comprehensive statistical table 4, it is shown that the average value of the six measurement data is 92.490, and the average correlation coefficient between the measurement data is 0.307, which indicates the partial description statistics of the measurement data as a whole, and can pass the average value, variance, etc. The information is further understood on the scale. The analysis result of the state after deleting a measurement data is shown in table 5. When the item A is deleted, the internal consistency of the evaluation scale becomes 0.750, which is obviously improved, and the correlation coefficient between the C item and other measurement data is the largest. For 0.839, the R^2 value reaches the highest at the same time, which is 0.963, which can be explained as 96.3% of the whole change, which can indicate that the consistency between the A measurement data and other measurement data is poor. From the variance analysis table 6, it can be concluded that under the analysis result of $F = 5.006$ and sig of 0.01, the overall evaluation of the evaluation scale is better, and the average average of the six measurements is 92.4897. The result is relatively stable, and Tukey can be used at the same time. After the additive test, it can be known that after the power conversion of the project, the evaluation scale will be more identifiable. In the correlation coefficient table 7, it is shown that the α reliability coefficient is 0.649, the significance level is 0.01, and when the α reliability coefficient is between 0.6 and 0.7, the reliability of the evaluation scale is better. The quality evaluation system has good consistency. Since the value of α reliability coefficient ranges from 0 to 1, the closer the value is to 1, the better the reliability of the evaluation system. Therefore, the evaluation system of the teaching quality of college students still needs to be further improved by certain technical means. Through the reliability analysis of the college students' teaching quality evaluation system, it can be found that the correlation coefficient is 0.655 when the significance is 0.05, and the α reliability coefficient is 0.649 when the clone Bach coefficient is detected. Therefore, the overall stability of the evaluation scale and the internal consistency of the evaluation results are good. The teaching status of teachers is basically at a stable level. The evaluation results are true and complete, with certain reliability and utilization value. On the other hand, through the analysis of the results, it is known that there is still room for further improvement and improvement in the evaluation scale, and the

stability and effectiveness of the evaluation system can be continuously improved to ensure that the results of student evaluation are trustworthy and thus improve teaching. Manage work and promote quality management of teaching.

5. Summary

The evaluation results of college students' teaching quality are affected by many factors. The behaviors and attitudes of students' evaluation of teaching, the unreasonable use of evaluation results, and the use of uniform teaching quality evaluation indicators in various disciplines will lead to deviations in evaluation results. In the management of teaching activities and teacher assessment, it is not entirely dependent on the results of student evaluation. To improve and improve the reliability and validity of the evaluation system of college students' teaching quality, and to improve the teaching effect of teachers, it is necessary to conduct in-depth analysis and research on the evaluation data of college students. Through data analysis, the problems and laws of the current evaluation system are found and feedback to teachers in a timely manner. And students; from the perspective of students, build a student-centered teaching quality evaluation system, strengthen publicity and education for students' evaluation work, and guide students to treat evaluation work with an objective and fair attitude; rational use of evaluation results Correctly handle the relationship between teacher assessment and student evaluation results. In summary, adhere to the "student-oriented, evaluation and promote teaching" teaching philosophy^[5], build a reasonable and complete evaluation system of teaching quality, continue to reform in practice, in-depth analysis of evaluation results, improve students' evaluation The reliability and validity can maximize the promotion of teaching quality, promote the development of colleges and universities, and cultivate high-quality talents for the society.

References

- [1] Dan Li, BingLiu, YingLin, etc. Reliability and Validity of Teaching Evaluation in Colleges and Universities[J]. Journal of Mudanjiang Medical University. 2015, 36(03):135.
- [2] Dan Li, BingLiu, GaochengSong, etc. Analysis on the Reliability and Validity of the Teaching Evaluation Index System for College Students- A Case Study of Mudanjiang Medical College[J]. China Medical Education Technology. 2016, 30(04):375-376.
- [3] Tao Sun, Pengdong Du. Application of statistical software SPSS in test paper retest reliability calculation[J]. Experimental Technology and Management. 2008, 25(03): 89-91.
- [4] Leilei Han, Yaoyu Zhou, Xiaobing Huang. Reliability and Validity Analysis of Teaching Evaluation Index System[J]. Higher Education Forum. 2012, 03(26):59-60.
- [5] Xin Yin, Xiaomei Zhang, Yimei Xie. The Status Quo, Problems and Countermeasures of College Students' Evaluation of Teaching — — Taking 20 "985" Universities as Examples[J]. Textile and Clothing Education. 2014, 8:287-290.