Evaluation of the experimental project performance assessment mode with the application of descriptive statistical analysis method

Renbao Liu
School of Public Administration, Shandong Technology and Business University, Yantai, China

Abstract: In order to solve the problem that currently, the performance assessment mode of most experimental projects only adopts the individual assessment while neglecting teamwork, puts forward the experimental project performance assessment mode based on team communication, which takes all the students participating in the project as a team, and divides the team into several groups according to project contents and tasks. Team members complete the project tasks jointly through team communication and group and individual quantitative scoring will be carried out according to the completion of tasks. After conducting the descriptive statistical analysis of an experimental project assessment performance of 223 students from 8 classes of 4 undergraduate majors in a university through Stata15.1, the results show that the results of the project assessment performance data all show a normal distribution by major, class, group and individual, the assessment results are ideal, and the assessment mode is feasible. The innovation is to use the Stata, and use the descriptive statistical analysis method to make an in-depth and careful statistical analysis of the experimental project performance assessment mode based on the team communication proposed. And the normality test is carried out to make an objective evaluation to the assessment results.

1. Introduction
The experimental project is an important part of the university curriculum, which combines the theory and practice of college students, improves their practical ability, and plays an important role in training their innovative spirit and innovative capabilities. [1] The experimental project performance assessment shall be fair and reasonable, and truly reflect the students’ experimental situation. At present, most of the experimental project performance assessment modes only use individual assessment methods, including the assessment in terms of attendance, project completion, answer of questions, quality of experimental report writing, etc. There are some shortcomings in this assessment mode. Firstly, it emphasizes the individual performance of the students while neglecting the teamwork. Secondly, the assessment cannot be accurate to everyone since there are so many students participating in the project, and the assessment is not objective and accurate enough. Thirdly, it pays attention to the assessment of results while neglecting the assessment of the project process. Fourthly, it mixes the assessment of the project itself with the auxiliary assessments relevant to the project, and weakens the role of the project. [2-5] Based on this, this paper proposes an experimental project performance assessment mode based on team communication, which separates students’ attendance,
answer of questions and evaluation of experimental reports as a separate assessment of individual students, while taking all the students participating in the project as a team, and dividing the team into several groups according to the project contents and tasks, and taking the group assessment performance as the common performance of each group members to achieve the team communication quantitative assessment based on groups to mobilize the enthusiasm and initiative of each student to participate in the project, enhance their team communication awareness, and improve their team cooperation capabilities. This paper takes the actual assessment results of an experimental course in a university as the sample, and uses the Stata, and uses the descriptive statistical analysis method to conduct detailed descriptive statistics and normality test to evaluate the feasibility of this mode. [6]

2. Experimental Principle and Subject

2.1. Experiment principle

2.1.1 Assessment mode.
The core of the experimental project performance assessment mode based on team communication is the quantitative assessment of team communication that takes the group as a unit, and takes all the students participating in the project as a team, and divides the team into several groups and a group leader is determined for each group. During the project process, group members jointly complete the project tasks through team communication. And quantitative scoring is carried out according to the completion of tasks, and group scoring is treated as the common score of the group members. The average score is taken as the final assessment performance of the individual experimental project (the group leader has a 5% bonus point).

2.1.2 Single project scoring rules.
(1) Percentage scoring. For a project with a clear number or score target, a team percentage scoring is adopted. And the basic scoring rule is the ratio of the group or team score to the target number or score multiplied by one hundred with integers are taken. Suppose the target quantity or score is \( m \), the group or team score is \( n \), and \( \text{int} \) means that the calculation result takes an integer, then the group score \( s \) is:

\[
s = \text{int}\left(\frac{n}{m} \times 100\right)
\]

(2) Assessment ranking scoring. For a project that is not suitable for percentage scoring, according to the number of groups divided, the ranking and score distribution are carried out according to the completion of the group's tasks, and the score distributions of the groups with different rankings are also different. Taking the division of 7 groups as an example, from the first place to the seventh place, the score is 100 to 60 respectively. Please refer to Table 1 for details.
Table 1. Group assessment ranking and score distribution table

| Ranking | 1      | 2      | 3      | 4      | 5      | 6      | 7      |
|---------|--------|--------|--------|--------|--------|--------|--------|
| Score   | 90-100 | 85-89  | 80-84  | 75-79  | 70-74  | 65-69  | 60-64  |

(3) Subjective scoring. For the project that has requirements on project process or results but the percentage scoring or assessment ranking scoring is not suitable for use, the subjective scoring is applied. And scoring is carried out to the completion quality or result of each group according to the project scoring standards, and the score value is 100 to 60 points. Taking the classification of 5 levels as an example, please refer to Table 2 for the group subjective scoring standard and score distribution.

Table 2. Group subjective scoring standard and score distribution table

| Standards          | Total conformity | Conformity | Average | More non-conformity | Nonconformity |
|--------------------|------------------|------------|---------|---------------------|---------------|
| Score              | 90-100           | 80-89      | 70-79   | 60-69               | 0-59          |

2.1.3 Single project scoring rules of groups.
If there is only one task in the experimental project, the applicable percentage scoring, assessment ranking scoring or subjective scoring rules will be adopted for group project scoring. If the experimental project has 2 or more tasks, then the average value of the scores of all tasks will be taken as the final score of the group project.

2.1.4 Final scoring rules of groups.
The average value of the scores of all the experimental projects of the group is taken as the final score of the group, namely, the individual assessment scores of all members of the group (the team leader has another 5% bonus point).

2.2. Experiment subject
The assessment mode has been applied in an experimental course of a university with 8 experimental projects, including 5 projects that participated in the assessment, and 223 students from 8 classes of 4 undergraduate majors participated in these projects.

3. Experiment and Discussion

3.1. Experiment
In the course experiment, the experimental projects 1, 2 and 3 are single task projects. And project 1 is scored according to the subjective scoring rules, project 2 is scored according to the percentage scoring rules, and project 3 is scored according to the rules of the assessment ranking scoring. Project 4 and project 5 both include two tasks, which are scored respectively according to percentage scoring and assessment ranking scoring, and the average value of the scores of two tasks is taken as the final score of the project. Projects 1, 3, 4, and 5 are scored in groups, with the group score being the individual score of the student, and project 2 is scored with the team as a unit, and with the team score being the project score of the group and the individual. Finally, the performance assessment results of each project are sorted out and obtained according to the professional grades, classes, groups and individuals.

3.2. Analysis and discussion
Using the Stata15.1 statistical analysis software, the statistical analysis is carried out according to the assessment performance of professional grades, classes, groups and individuals.
3.2.1 Statistical analysis based on professional grade assessment performance.

Detailed descriptive statistics is carried out to experimental projects 1 to 5 (p1, p2, p3, p4, p5) and average scores (as), and the results are shown in Table 3.

Table 3. Detailed descriptive statistical analysis results based on each project assessment performance of professional grades

| Variable | Obs | Mean  | Std. Dev. | Variance | Min  | Max  | 50%  | Skewness | Kurtosis |
|----------|-----|-------|-----------|----------|------|------|------|----------|----------|
| p1       | 4   | 82.08 | 2.29      | 5.24     | 79   | 84   | 83   | -0.78    | 1.97     |
| p2       | 4   | 70.25 | 17.63     | 310.92   | 49   | 92   | 70   | 0.05     | 1.97     |
| p3       | 4   | 75.75 | 1.37      | 1.88     | 75   | 78   | 75   | 1.15     | 2.33     |
| p4       | 4   | 74.88 | 2.82      | 7.95     | 71   | 77   | 76   | -1.02    | 2.21     |
| p5       | 4   | 74.79 | 2.61      | 6.80     | 71   | 78   | 75   | -0.48    | 2.03     |
| as       | 4   | 76.55 | 4.32      | 18.63    | 71   | 82   | 76   | 0.08     | 2.00     |

It can be seen from the above table that the average value of p1 to p5 is between 70.25 and 82.08; the standard deviation is between 1.37 and 17.63; the variance is between 1.88 and 310.92; the minimum value is between 49 and 79; the maximum value is between 77 and 92; the median is between 70 and 83; the skewness is between -1.02 and 1.15, basically around 0; and the kurtosis is between 1.97 and 2.33, both less than 3, but close to 3. On the whole, there is no obvious abnormal data, the data are normally distributed, and the data quality is high.

3.2.2 Statistical analysis based on class assessment performance.

(1) Descriptive statistical analysis of data. After carrying out descriptive statistical analysis of data to the projects 1 to 5 (p1, p2, p3, p4, p5) and the average scores (as), the results are shown in Table 4.

Table 4. Detailed descriptive statistical analysis results based on each project assessment performance of classes

| Variable | Obs | Mean  | Std. Dev. | Variance | Min  | Max  | 50%  | Skewness | Kurtosis |
|----------|-----|-------|-----------|----------|------|------|------|----------|----------|
| p1       | 8   | 82.14 | 3.21      | 10.33    | 76   | 85   | 82   | -0.92    | 2.92     |
| p2       | 8   | 70.25 | 16.32     | 266.50   | 49   | 92   | 70   | 0.05     | 1.97     |
| p3       | 8   | 75.48 | 4.17      | 17.42    | 69   | 80   | 77   | -0.57    | 1.75     |
| p4       | 8   | 74.62 | 5.87      | 34.50    | 67   | 84   | 73   | 0.53     | 2.01     |
| p5       | 8   | 74.70 | 4.58      | 21.00    | 68   | 82   | 75   | 0.01     | 2.01     |
| as       | 8   | 76.43 | 4.38      | 19.21    | 71   | 85   | 76   | 0.60     | 3.06     |

It can be seen from the above table that the average value of p1 to p5 is between 70.25 and 82.14; the standard deviation is between 3.21 and 16.32; the variance is between 10.33 and 266.50; the minimum value is between 49 and 76; the maximum value is between 80 and 92; the median is between 70 and 83; the skewness is between -0.92 and 0.60, basically around 0; and the kurtosis is between 1.75 and 3.06, basically around 3. On the whole, there is no obvious abnormal data, and the data quality is high.

(2) Normality test of data. The normality test is carried out to experimental projects 1 to 5 (p1, p2, p3, p4, p5) and average scores (as), and the results are shown in Table 5.

Table 5. Normality test results based on each project assessment performance of classes

| Variable | Obs | Pr(Skewness) | Pr(Kurtosis) | adj chi2(2) | Prob>chi2 |
|----------|-----|--------------|--------------|-------------|-----------|
| p1       | 8   | 0.12         | 0.35         | 3.69        | 0.16      |
| P2       | 8   | 0.94         | 0.67         | 0.19        | 0.91      |
| P3       | 8   | 0.34         | 0.40         | 1.93        | 0.38      |
| P4       | 8   | 0.37         | 0.73         | 1.01        | 0.60      |
| P5       | 8   | 0.98         | 0.73         | 0.12        | 0.94      |
| as       | 8   | 0.31         | 0.28         | 2.68        | 0.26      |

It can be seen from the above table that the test results of the skewness and kurtosis of all the experimental projects and their average scores are greater than 0.05, accepting the null hypothesis of the normal distribution. Considering the skewness in combination with the kurtosis, on the whole, the
Prob>chi2 values of all test results are greater than 0.05, which all accept the null hypothesis of normal distribution, and the data show a significant normal distribution.

To summarize, there is no obvious abnormality in the class assessment performance data, the data show a significant normal distribution. The data quality is good, and the assessment mode is feasible.

3.2.3 **Statistical analysis based on the group assessment performance.**

(1) Descriptive statistical analysis of data. Detailed descriptive statistics is carried out to experimental projects 1 to 5 (p1, p2, p3, p4, p5) and average scores (as), and the results are shown in Table 6.

**Table 6.** Detailed descriptive statistical analysis results based on each project assessment performance of groups

| Variable | Obs | Mean  | Std. Dev. | Variance | Min | Max | 50% | Skewness | Kurtosis |
|----------|-----|-------|-----------|----------|-----|-----|-----|----------|----------|
| p1       | 28  | 81.96 | 7.12      | 50.63    | 70  | 95  | 83  | -0.25    | 2.15     |
| p2       | 28  | 70.25 | 15.55     | 241.82   | 49  | 92  | 70  | 0.05     | 1.97     |
| p3       | 28  | 75.89 | 9.53      | 90.84    | 60  | 90  | 75  | -0.03    | 1.82     |
| p4       | 28  | 74.75 | 12.53     | 156.99   | 50  | 95  | 77  | -0.25    | 2.03     |
| p5       | 28  | 74.71 | 11.23     | 126.19   | 50  | 98  | 75  | -0.02    | 2.71     |
| as       | 28  | 76.52 | 6.03      | 36.35    | 63  | 92  | 76  | 0.23     | 3.05     |

It can be seen from the above table that the average value of p1 to p5 is between 70.25 and 81.96; the standard deviation is between 6.03 and 15.55; the variance is between 36.35 and 241.82; the minimum value is between 49 and 70; the maximum value is between 90 and 98; the median is between 70 and 83; the skewness is between -0.25 and 0.23, basically around 0; and the kurtosis is between 1.82 and 3.05, basically around 3. On the whole, there is no obvious abnormal data, and the data quality is high.

(2) Normality test of data. The normality test is carried out to experimental projects 1 to 5 (p1, p2, p3, p4, p5) and average scores (as), and the results are shown in Table 7.

**Table 7.** Normality test results based on each project assessment performance of groups

| Variable | Obs | Pr(Skewness) | Pr(Kurtosis) | adj chi2(2) | Prob>chi2 |
|----------|-----|--------------|--------------|-------------|-----------|
| p1       | 28  | 0.53         | 0.30         | 1.62        | 0.44      |
| p2       | 28  | 0.90         | 0.11         | 2.79        | 0.25      |
| p3       | 28  | 0.94         | 0.03         | 4.60        | 0.10      |
| p4       | 28  | 0.53         | 0.16         | 2.57        | 0.28      |
| p5       | 28  | 0.96         | 0.91         | 0.01        | 0.99      |
| as       | 28  | 0.57         | 0.55         | 0.72        | 0.70      |

It can be seen from the above table that only the kurtosis of experimental project 3 is 0.03, less than 0.05, rejecting the null hypothesis of normal distribution, and other test results of experimental projects 1-5 and their average scores are greater than 0.05, accepting the null hypothesis of normal distribution. Considering the skewness in combination with the kurtosis, on the whole, the Prob>chi2 values of all test results are greater than 0.05, which all accept the null hypothesis of normal distribution, and the data show a significant normal distribution.

To summarize, there is no obvious abnormality in the group assessment performance data, the data show a significant normal distribution. The data quality is good, and the assessment mode is feasible.

3.2.4 **Statistical analysis based on individual assessment performance.**

(1) Descriptive statistical analysis of data. Detailed descriptive statistics is carried out to experimental projects 1 to 5 (p1, p2, p3, p4, p5) and average scores (as), and the results are shown in Table 8.
Table 8. Detailed descriptive statistical analysis results based on each project assessment performance of individuals

| Variable | Obs | Mean | Std. Dev. | Variance | Min | Max | 50% | Skewness | Kurtosis |
|----------|-----|------|-----------|----------|-----|-----|-----|----------|----------|
| p1       | 223 | 82.09| 6.91      | 47.77    | 70  | 95  | 80  | -0.24    | 2.18     |
| p2       | 223 | 68.77| 15.31     | 234.43   | 49  | 92  | 68  | 0.15     | 1.98     |
| p3       | 223 | 75.76| 9.35      | 87.48    | 60  | 90  | 75  | -0.01    | 1.82     |
| p4       | 223 | 74.81| 12.18     | 148.45   | 50  | 95  | 77  | -0.24    | 2.02     |
| p5       | 223 | 74.57| 10.75     | 115.49   | 50  | 98  | 73  | 0.04     | 2.68     |
| as       | 223 | 76.20| 6.06      | 36.73    | 63  | 96  | 76  | 0.27     | 3.11     |

It can be seen from the above table that the average value of p1 to p5 is between 68.77 and 82.09; the standard deviation is between 6.06 and 15.31; the variance is between 36.73 and 234.43; the minimum value is between 49 and 70; the maximum value is between 90 and 98; the median is between 68 and 80; the skewness is between -0.24 and 0.27 basically around 0; and the kurtosis is between 1.82 and 3.11, basically around 3. On the whole, there is no obvious abnormal data, and the data quality is high.

(2) Normality test of data. The normality test is carried out to experimental projects 1 to 5 (p1, p2, p3, p4, p5) and average scores (as), and the results are shown in Table 9.

Table 9. Normality test results based on each project assessment performance of individuals

| Variable | Obs | Pr(Skewness) | Pr(Kurtosis) | adj chi2(2) | Prob>chi2 |
|----------|-----|--------------|--------------|-------------|-----------|
| p1       | 223 | 0.13         | 0.00         | 15.95       | 0.00      |
| p2       | 223 | 0.34         | 0.00         | 33.71       | 0.00      |
| p3       | 223 | 0.97         | 0.00         | 70.25       | 0.00      |
| p4       | 223 | 0.13         | 0.00         | 29.05       | 0.00      |
| p5       | 223 | 0.79         | 0.34         | 1.01        | 0.60      |
| as       | 223 | 0.10         | 0.58         | 3.02        | 0.22      |

It can be seen from the above table that the test values of the skewness of the experimental projects 1, 2, 3 and 4 are all greater than 0.05, accepting the null hypothesis of the normal distribution; the kurtosis is 0.00, less than 0.05, rejecting the null hypothesis of normal distribution. Considering the skewness in combination with the kurtosis, on the whole, the Prob>chi2 values of the test results from projects 1 to 4 are all less than 0.05, rejecting the null hypothesis of normal distribution. The skewness, kurtosis and the overall test value of the experimental project 5 and the average score are both greater than 0.05, accepting the null hypothesis of normal distribution. Project 5 and individual average score data all show a significant normal distribution.

On the whole, since the average score is the final assessment result of individual scores, the detailed descriptive statistics and normality test results of the average scores of individual assessment performance show that there is no obvious abnormality in the average scores of individual assessment performance, and the data show a significant normal distribution. The data quality is good, and the assessment mode is feasible.
4. Conclusions
The experimental project performance assessment mode based on team communication takes professional grades as a team, and groups in the team as the unit to complete each experimental project through communication and coordination of each group. According to the task completion of the team or groups, it adopts quantitative scoring, ranking scoring or subjective scoring and other scoring rules, and takes the project scores of the team or groups as the assessment performance of each member of the group, and takes the average score of all projects as the final assessment performance of each student (the group leader has a 5% bonus point). Judging from the actual assessment results, the main performance data all show normal distribution, and this assessment mode is practically feasible, with promotion and application value.

The shortcomings are that the experimental data in this paper come from 8 classes of 4 majors in one grade of an experimental course. The overall sample size is not enough, which may have some impact on the experimental results. In its application in the following professional grades of this experimental course and other experimental courses, the sample size of this assessment mode will be continuously increased to further improve the data quality.

Some statistical indicators in the sample data are not significant. For example, the group's performance kurtosis of project 2 and the individual performance kurtosis of projects 1-4 show a non-normal distribution, indicating that the capabilities of different majors or student groups in some projects are objectively quite different, and the distribution of people with different scores is quite different. This phenomenon may change with the increase of the sample size, but it may also be an objective phenomenon of some majors or student groups, which will not change with the change of sample sizes.

In summary, judging from the experimental results, the experimental project performance assessment mode based on team communication is feasible and worthy of application in the experimental courses suitable for grouping. In future application, the application results of this mode will be further tested through more abundant experimental data.

References
[1] Feng Xu, Yong Sun, Hongxi Zhang and Shuo Ji 2019 Exploration and practice of experimental teaching model in accord with requirements of applied talent training Experimental Technology and Management vol 36 (06) p 214-215+224. (In Chinese)
[2] Nan Shi, Xianhui Zhu, Pengtao Yuan, Caifeng Ma and Lin Sang 2018 Research on the Experimental Teaching Mode under the Target of Applied Talents Training Heilongjiang Education (Theory and Practice) (04) p 44-46. (In Chinese)
[3] Hua Qiang, Zhengwang Li and Shihui Wu 2018 Exploration of practical teaching assessment method in colleges and universities Experimental Technology and Management vol 35 (06) p 170-173. (In Chinese)
[4] Yancheng Yuan, Nianling Lv, Lijuan Lai and Ruixiang Yin 2019 Scientific Design and Process Management Scheme of Experimenta Experimental Technology and Management vol 36 (01) p 192-198+216. (In Chinese)
[5] Zhifang Yang, Zehui Chen, Xin Min, Anlin Chen, Zeling Dong, Xianlian Chen, Huan Yang and Ting Zhou 2017 Establishment and Implementation of Formative Evaluation Model of Clinical Microbiology Testing Technology Experiment Course Continuing Medical Education vol 31 (11) p 56-58. (In Chinese)
[6] Xihua Chen, Haining Huang, Peijie Huang 2019 An Evaluation of the Scientifcificity of Test Papers Based on Students’ Score Analysis - Application of Data Mining and Feature Description Journal of Beijing City University (01) p 38-43. (In Chinese)