

Assessment by Psychophysical Methods for Design Courses of Optical Discipline

Weige Lv, Xiaodong Zheng, Xiaotong Li, Lingying Jiang, Zhaofeng Cen
College of Optical Science and Engineering, Zhejiang University

ABSTRACT

In the teaching of design courses, the process of students to complete the designs is considered more important than the final results, and usually there are no standard answers for these designs. Thus, the research on the process evaluation method in design courses is of great significance. Taking the assessments of two design courses in optical discipline as examples, the psychophysical experimental methods are introduced into the process evaluation of the design courses, i.e. the process evaluation given by each student to others in forms of classification, sorting or grading are adopted as one trial of a psychophysical experiment, which generate many experimental data of mutual evaluation. Based on these data, evaluation results are measured scientifically using the statistical method. Furthermore, through correlation analysis and regression analysis of these data, the relationship among various aspects of different process can be studied. It is easier for students to understand and accept their assessment results, and more helpful for teachers to analyze the influencing factors in their teaching process.

Keywords: psychophysical methods, process evaluation, design courses

1. INTRODUCTION

In recent years, teaching reform in colleges and universities of China pay more and more attentions on three development transformation: change from Teaching-centered Model to Learning-centered Model, change from classroom teaching to the combination of inside and outside class, change from outcome evaluation to process evaluation^{1,2}. The transformation of evaluation methods is an important condition to ensure the first two changes.

According to the “developing evaluation” theory put forward by the British scholar Latoner and Crift in the end of the 20th century, the process of evaluation should be people-oriented and pluralistic-orientated. Then students are encouraged to participate in the evaluation process, which will promote students' learning and development through integrating the teaching process and the evaluation process.

In this paper, the psychophysical method is introduced into the process evaluation of design courses which usually have multiple evaluation indexes because of their comprehensiveness^{3,4}. Taking two design courses Design of optical systems and Design of opto-mechanism structures as examples, students' assessments in forms of classification, sorting or grading are adopted as one trial of a psychophysical experiment. Based on these experimental data, evaluation results are measured scientifically using statistic methods. Furthermore, through correlation analysis and regression analysis of these data, the relationship among various aspects of different process can be studied.

2. METHODS

2.1 Experimental Design

In psychophysics experiments, the experimental condition, methods and procedure should be designed carefully. For design courses, usually, all the students in the course are taken as the observers or evaluators, and the design works are shown to all the observers in the form of PPT presentation within several minutes, as well as some material such as videos, documents or programs. The evaluation indexes are listed, and the instruction about how to evaluate them are given to the observers before the experiment. There are two methods to determine the indexes. One method is to divide the design course into different stages, each stage corresponding to an index. Another method is to extract the indexes according to the training targets of the courses.

Then the students give their evaluations individually or in groups. According to the characteristics of the evaluation sub

*lvweige@zju.edu.cn; phone 086 0571 87951681

item, different psychophysical experimental methods are chosen. For some concrete index which can be easily judged, the method of magnitude estimation is often adopted to obtain the scores directly, i.e. the observers are asked to assign numbers in proportion to the magnitude of the stimulus. If some evaluation index is abstract and the evaluation is difficult to score directly, the category judgment method can be adopted. For example, observers are asked to rate an index or an attribute using a 7-point verbally-labelled category scale, with '7' corresponding to the highest level, '4' to the average level and '1' to the lowest level. The category numbers assigned by each observer are converted into equal-interval scale values through Case V of Thurstone's law of comparative judgments⁵.

In the case of the design of optical system, the evaluation experiment was carried out in class with all the students as observers to investigate the relationship between the total design quality and its four aspects: the aberration calculation program, the optical lens design, the aberration optimization and the design drawings, and the total quality. The assessment was made in groups, where the evaluating grouping is consistent with the design grouping. Instructions for evaluating are explained to them by teachers before experiments. Therefore, a training session was not given to them. All the designed works were used as stimuli in the assessments, and were shown to observers in random sequence with 10 minutes PPT representation and 5 minutes question. Each work was assessed by N groups of observers using magnitude estimation method. Thus, the experiment was divided into N observing sessions. Each session contained 5N assessments (1 work \times 5 attributes \times N observer groups) and lasted for approximately 15 minutes.

2.2 Data Analysis

This section describes statistical measures that are used to analyze the psychophysical experimental data and to develop models in this research. The following measures can be calculated through EXCEL or MATLAB program.

2.2.1 Coefficient of Variation

The coefficient of variation, CV, which is a statistical measure to represent the agreement between two sets of data, expresses standard deviation as a percentage of the mean. In psychophysical experiments, observer variations are usually computed using this statistical measure. The CV value is limited to 0-100, and the greater the CV value, the greater the deviation of the two sets of data; the CV value is 0, indicating that the two sets of data are the same. Therefore, the CV value can be used to measure the data stability between the observers, that is, the inter observer accuracy.

2.2.2 Coefficient of Correlation

In the mathematical statistics, when the two variables X and Y are normal continuous variables, and satisfy the linear relationship between the two, Pearson product-moment correlation coefficient, which is defined as the quotient of covariance and standard deviation, can be adopted to measure the linear correlation between these two variables. The value of the correlation coefficient always lies in the range -1 to 1 . If the correlation between two sets of data is positive and close to 1 , it can be said that the two sets of data have a strong positive linear correlation.

2.2.3 Analysis of Variance, ANOVA

To study the influence of experimental conditions or physical factors, the mathematical method of single factor analysis of variance (ANOVA analysis) can be used. The principle is: the general variable due to a variety of factors will show fluctuations in the data, for which there are two sources: the variations within the group caused by uncontrollable systematic errors or random factors of chance, and the differences between the groups caused by the controllable experimental conditions. The test value of ANOVA is defined as the ratio F between the inter and intra group differences. The F value should be compared with the critical value F_c , and if $F > F_c$, significant difference can be observed between the two groups.

2.2.4 Multiple linear regression

Under some assumptions on the regression function and errors terms, using the approach of multiple linear regression, the relationship between a scalar dependent variable y and multiple explanatory variables denoted X can be modeled, with the dependent variable y expressed as a linear combination of multiple explanatory variables. The multiple linear regression model can not only find out the main factors affecting the dependent variables, but also explain the relationships between them, and make predictions for the values of dependent variable when explanatory variables are given new values.

3. RESULTS

Take the evaluation of the course 'Design of the opto-mechanical structure' as an example, the experimental data were analyzed. The evaluation experiment was carried out in class with all the 62 students in 21 groups as observers to assign numbers in proportion to the magnitude of the total design quality and its six attributes or indexes, which are listed in the first column of Table 1. The assessment was made in groups for all the design works. No training sessions but concrete evaluation instructions were given to them. Each design work was exhibited to observers in random sequence with 10 minutes of PPT representations and 5 minutes of questions. The experiment was divided into 21 observing sessions. Each session contained 7*21 assessments (1 work × 7 attributes × 21 observer groups) and lasted for approximately 15 minutes.

3.1 Observer accuracy

The CV values between the individual observer data and mean data of all observers for the 7 evaluation indexes in this experiment are calculated as the inter-observer accuracy. Table 1 summarizes the resulting CV values for inter-observer agreement in terms of mean, maximum and minimum. Most CV values are less than 20, suggesting good consistency between the observers.

Table 1 Inter observer accuracy for different evaluation indexes

Evaluation index	Mean	Max	Min
Overall Quality	8.6	13.3	6.2
Design Method	10.2	16.2	3.9
Design Index	15.7	22.5	9.3
Modeling	8.7	17.7	5.1
Feasibility	16.8	22.2	9.5
Cooperation	12.5	16.7	9.1
Presentation	12.3	16.8	8.9
Mean	12.1	17.9	7.4

3.2 Relationships between different evaluation indexes

To investigate the relationships between different evaluation indexes, Pearson correlation coefficients are computed between the evaluation data sets for different indexes, as shown in Table 2.

Table 2 Pearson correlation coefficients between the evaluation data set for different indexes

Evaluation index	Overall Quality	Design Method	Design Index	Modeling	Feasibility	Cooperation	Presentation
Overall Quality	1.000						
Design Method	0.836	1.000					
Design Index	0.756	0.550	1.000				
Modeling	0.861	0.685	0.525	1.000			
Feasibility	0.621	0.411	0.581	0.324	1.000		
Cooperation	0.634	0.446	0.431	0.382	0.411	1.000	
Presentation	0.682	0.508	0.446	0.475	0.335	0.580	1.000

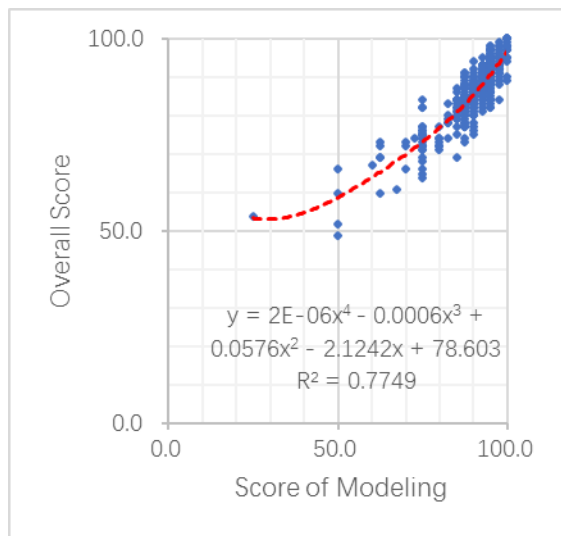
In Table 1, the Pearson correlation coefficients more than 0.5 are labeled in bold. All the correlation coefficients between the evaluation data of Overall quality and each attribute or sub-index have values more than 0.6, indicating strong correlations between them. Specifically, the evaluation of Modeling has a very strong correlation to Overall quality evaluation with a correlation coefficient of 0.861, as well as Design method with a coefficient of 0.836. Next one

is Design index, then Presentation, Cooperation, and Feasibility in turn. High correlation of Design method and Modeling to Overall quality, exhibit that the main goal of our teaching which is to enable students to master the basic design method and modeling method of opto-mechanical structure has been recognized widely. For the combination of sub-indexes, moderate correlations exist between 6 pairs of them with the coefficients higher than 0.5: Design method and Modeling (0.685), Design method and Design index (0.550), Design method and presentation (0.508), Design index and Feasibility (0.581), Design index and Modeling (0.525), Team cooperation and Presentation (0.580). The correlation coefficient of other combinations is less than 0.5, which can be regarded as basically no correlation. The dependency of one to other sub-indexes exhibits obviously for some abstract indexes. For example, students are more inclined to compare the ideas, implementation and expression of the design method from the aspects of Design Index, Modeling and Presentation.

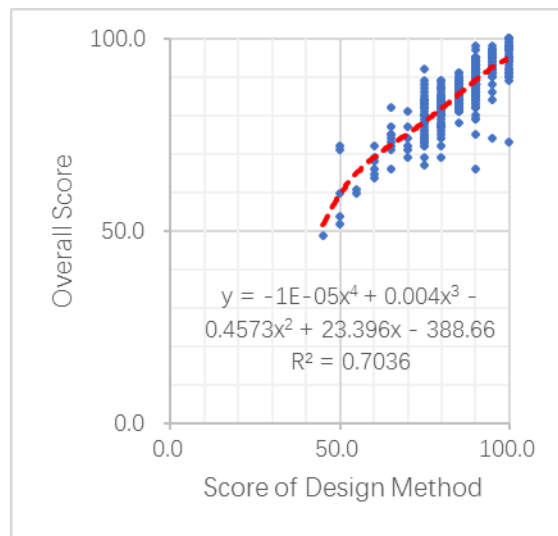
To furtherly examine the psychophysical relationships between Overall quality and each of its constituent sub-index, the scores of Overall Quality are plotted in Figure 1 (a) to (f) against those of Modeling, Design Method, Design Index, Presentation, Cooperation, and Feasibility respectively. A best-fit curve was also given so as to indicate the trend in the relationship between them.

Figure 1 (a) shows that the Overall score firstly increases slowly with an increase in Modeling score, then increases faster when Modeling score reaches some level. While in Figure 2 (b), a complementary trend is shown, i.e., the Overall score increases fast with an increase in Design Method Score at its low level, then increases slowly at its higher level. This indicates that, the higher the overall quality is, the more important role the modeling plays. For poor design works, they usually stayed at the initial stages of concept design which can be evaluated more effectively by their design methods but not modeling works that they didn't include.

Figure 1 (c) to (f) shown very discrete distribution of evaluation data, which indicates that students might use different criteria in their evaluation of these indexes. It's suggested that more clear instruction about how to evaluate these indexes given to students before evaluation.



(a)



(b)

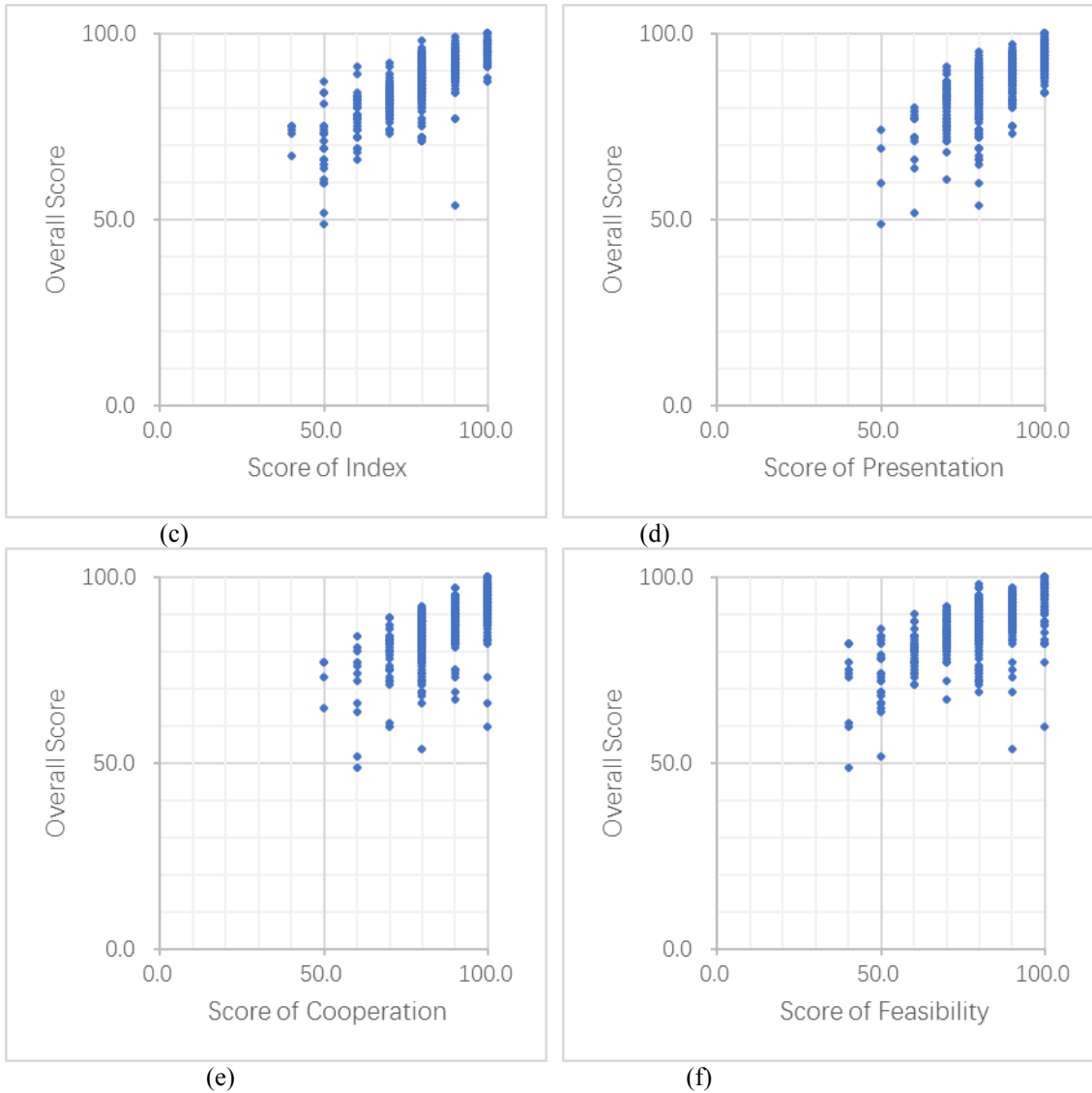


Figure 1 The psychophysical relationships between Overall quality and each of its constituent sub-index

3.3 Multiple regression representation of overall scores

It's hoped that, the overall evaluation of the design course can be obtained through evaluation of each important index. For this purpose, the appropriate key evaluation items should be tested and selected, and combined with different weights to describe the overall evaluation. In the modeling phase, the mathematical method of multiple linear regression is used.

According to the result of ANOVA using different design works as the single factor, the evaluation of cooperation is found change little across design works. In fact, the team with poor cooperation but outstanding team members might design works of high level. Thus, cooperation is suitable to be assessed within group members according to their contributions, and not included in the representation of the overall quality of a group work. Another try is to remove the abstract index such as Design Method, which may be replaced by the combination of other related indexes. Three kinds of representations of Overall Quality using different combinations of indexes are listed in Table 3. Teachers can set the weight values of different indexes according to the coefficients.

Table 3 Multiple regression representation of overall score

Evaluation Index	Intercept	Design Method	Design Index	Modeling	Feasibility	Cooperation	Presentation	Adjusted R Square
Coefficients	0	0.202	0.104	0.397	0.101	0.094	0.102	0.9977
Coefficients	0	0.219	0.108	0.413	0.117	-	0.147	0.9976
Coefficients	0	-	0.131	0.549	0.135	-	0.185	0.9973

4. CONCLUSION

Using the psychophysical experimental method and the corresponding statistic method, the relationships between different evaluation indexes of design works in the course “Design of opto-mechanistic structure” are analyzed and discussed, and the multiple regression representation of overall quality is given in the form of linear combination of weighted sub-indexes. This provides a scientific method for the process evaluation in design courses. It’s easier for students to understand and accept their assessment results, and more helpful for teachers to analyze the influencing factors in their teaching process.

REFERENCES

- [1] Lu guodong, He qinming, Zhang cong, [Reform of teaching method: focusing on process and interaction: best practices at Zhejiang university (in Chinese)], Zhejiang University Publishers, Hangzhou, 1~22(2013)
- [2] Zheng xiaodong, Wen chunao, Wang Xiaoping, et al. "Investigation in Grading System of Optical Engineering Lab Courses in World Famous Universities (in Chinese)," Research and Exploration in Laboratory 30(7),115~117(2011)
- [3] Lv weige, Zheng xiaodong, et al. "The application of ‘Questionnaire Star’ software platform in teaching process evaluation[C](in Chinese)," Proceedings of the Conference of Chinese Optical Society Sichuan, China (2015)
- [4] J. A. Schinka, W. F. Velicer, and I. B. Weiner. [Handbook of Psychology: Research Methods in Psychology] 2nd Edition. John Wiley&Sons, Hoboken, New Jersey, United States, (2003)
- [5] L. L. Thurstone "A law of comparative judgment," Psychology Review 34, 273-286(192)