I.J. Education and Management Engineering **2012, 3, 39-45**
Published Online March 2012 in MECS (http://www.mecs-press.net)
DOI: 10.5815/ijeme.2012.03.06

Available online at http://www.mecs-press.net/ijeme

Exploration and Practice for Evaluation of Teaching Methods①

Xiaodong Wang[a], Yingjie Wu[b]

[a] *College of Mathematics and Computer Science, Quanzhou Normal University, Quanzhou 362000, China*
[b] *College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350002, China*

Abstract

This paper discusses one of a series of achievements of teaching reform for the national excellent course "Algorithms and Data Structures". One of the most valuable and most difficult reforms is to strengthen the practical aspects of the curriculum, while strengthening the teacher-student interactive teaching. Another innovation closely related to this teaching reform is a comprehensive scientific evaluation of teaching methods for this complex teaching process. This evaluation method will evaluate the whole process of teaching by a dynamic multi-factor evaluation. It will effectively promote the teaching level continues to increase.

Index Terms: Evaluation; Interaction; Data Structures

1. Introduction

We have conducted a series of teaching reform for "Algorithms and Data Structures", the advanced course of computer science [1]. One of the most important and most difficult of the reform is to strengthen the practical aspects of the curriculum, while strengthening the teacher-student interactive teaching. The reform also brings up a problem how to accurately evaluate the effectiveness of teaching and teaching process and student's learning outcomes.

The traditional approach of evaluation of teaching is through the course examination. The states of the student's mastering of the basic concepts, basic algorithms and basic techniques are evaluated by the course examination. Each student's test scores to evaluate their learning. Test scores of all students to evaluate teaching effectiveness.

This evaluation method of teaching has its rationality. It has been used in our teaching for many years, and generally accepted by the majority of teachers and educational administrators. However, this teaching evaluation method looks a little outdated for our current education reform. It can not evaluate the entire of our

① This work was supported by the Natural Science Foundation of Fujian under Grant No.2009J01295 and the Haixi Project of Fujian under Grant No.A099.

dynamic teaching process completely and accurately. So it is impossible to make a comprehensive evaluation of all aspects of teaching using this evaluation method.

Therefore, how to scientifically evaluate the teaching process and teaching effectiveness of the course "Algorithms and Data Structures" becomes a new topic for our teaching reform. It involves many aspects of teaching reform, including the innovation of teacher's teaching ideas and teaching styles and students' studying approaches also involves teaching management model. A scientific teaching evaluation should include evaluation in the field of learning and emotion and the field of operation.

The main basis for the evaluation in the field of learning includes curriculum and teaching materials, course exams questions and the validity and reliability of questions, which is the traditional final examination of course. While the evaluations in the field of learning and in the field of operation are often difficult to grasp and the most easily be overlooked aspects. These evaluations must include the aspects of students' self-learning ability, basic skills, learning attitude and perseverance, independent thinking, creative thinking, learning autonomy, cooperation, attitude and sense of competition and so on. These aspects are the basic elements of quality education.

The interaction and practice teaching reforms of the course "Algorithms and Data Structures" precisely reflect these aspects of teaching effectiveness. It is impossible to evaluate these new and important aspects in the field of learning by using traditional evaluation methods. Therefore, we must break through the traditional method of teaching evaluation and create a new comprehensive teaching evaluation method to evaluate the teaching process and practice of the course completely and accurately.

In the following 3 sections we describe our comprehensive scientific evaluation of teaching methods for the complex teaching process and teaching reforms of the course "Algorithms and Data Structures".

In section 2 we describe our new comprehensive teaching evaluation methods.

In section 3 we discuss the comprehensive evaluation effectiveness of the new method.

Some concluding remarks are in section 4.

2. New Comprehensive Evaluation Methods

2.1. Selection of Teaching Evaluation Indicators

In order to explore a new comprehensive teaching evaluation method, first of all it is reasonable to consider how to set up observation and evaluation points and the assessment indicators [2]. On one hand we should focus on the important functions of the comprehensive teaching evaluation method in induction, summarize, synthesize, and sublimation of knowledge, on the other hand we can not ignore the formation and continuity of the teaching evaluation process. Therefore, we made the following comprehensive reform for evaluation contents and forms and observation points.

The first is the diversity of the evaluation forms. We arrange a midterm exam and a final exam in one semester. Each test using a teaching evaluation form combining a theory test and an open-book practice exam. This new teaching evaluation method has changed the former evaluation form. The new added exam for students' practical skills assessment can evaluate students' theory, abstraction and design capabilities. Besides the midterm and final exams, we have arranged a practice exercises after each teaching unit. Students submit their programming codes online and take part in the excellent homework assignments for the practice exercises. This evaluation method can reflect the dynamic process of gradual teaching, avoid the problem of too few observation points in the midterm and final exams to cover all the teaching content of the course, tend to be more scientific and more fair.

Second is the evaluation content diversity. We evaluate the students' course studies from several aspects such as their lecture attendance, finish the homework independently and their learning attitude, excellent homework communication in class, answer questions, actively participate in interactive teaching, etc. In order to evaluate the students' homework fairly, the curriculum group developed an automatic assessment system, which can perform an automatic laboratory-skill assessment for students' exercises.

Table 1. Evaluation Indices of the course

Name	Meaning	Score
X_1	Practice1	0-12
X_2	Practice 2	0-12
X_3	Practice 3	0-12
X_4	Practice 4	0-12
X_5	Practice 5	0-12
X_6	Practice 6	0-24
X_7	Practice 7	0-24
X_8	Practice 8	0-24
X_9	Practice 9	0-12
X_{10}	Practice 10	0-12
X_{11}	Midterm practice exam	0-500
X_{12}	Final theory exam	0-100
X_{13}	Final practice exam	0-150
X_{14}	Attendance	0-100
X_{15}	Independence	0-100
X_{16}	excellent homework	0-100

Third is diversification of the evaluation. In the traditional teaching evaluation, students are always evaluation objects and in a passive position. In the teaching process of our course "Algorithms and Data Structures", the subjects of evaluation not only includes teacher but also includes the students' self assessment and evaluation of interaction between students. This method arouses the students' active participation enthusiasm and initiative.

The roles of students changed from pure evaluation object into a subject of evaluation. The combination of the evaluation subject and the evaluation object makes the evaluation process an equal exchange process between teachers and students, which can provide us multi-angle, multi-layered evaluation information.

Based on the above consideration, we propose 16 evaluation indices for the course "Algorithms and Data Structures" as shown in table 1.

2.2. Weights of Teaching Evaluation Indicators

In the evaluation of teaching effects of the course "Algorithms and Data Structures" using the 16 indicators described above we should also consider the impacts of the indicators, which are the weights of indicators in teaching evaluation. Statistical method is the commonly used method for determining the indicator weights. Since there are correlations among the 16 indicators, we can use principal component analysis method to the 16 indicators to get the load of principal component in each indicator. Then the load of the first principal component in each indicator can be used as the weight of each indicator.

2.3. Composite Scores of the Course Grade

We get the weights of the 16 indicators in the comprehensive evaluation by the principal component analysis method and use the weights to construct a comprehensive indicator of scores.

The 16 values of indicator for each student are substituted into the comprehensive indicator and then the students' comprehensive achievements can be computed accordingly.

Let the comprehensive scores of *n* students be $f_1, f_2, \ldots, f_n$.

In order to facilitate students' grades we can transform the comprehensive scores of students into 100 points. The transformation steps are as follows:

(1) Sort $f_1, f_2, \ldots, f_n$ from low to high; Denote the *n* sorted comprehensive scores by $F_1, F_2, \ldots, F_n$.

According to the students' learning state in each class, teachers predetermine the following index values: Failure rate *p* of students; Minimum final score *min* and Maximum final score *max*.

(2) Convert the comprehensive scores into 100 points using linear function conversion.

3. Evaluation Effectiveness of the New Method

We have conducted a comprehensive evaluation of teaching on the course "Algorithms and Data Structures" for the students majored in Computer Science and Technology in Fuzhou University using above method. There were total 327 students with 16 values of index scores for each student. For the correlation coefficient matrix of Table 3, we computed its eigenvalues and eigenvectors and obtained the first principal component factor loading by using our computer software. The combined load factor of 16 indicators is 3.81. The weight of each index is its index value divided by the load factor of 3.81. The results are shown in Table 2.

Table 3. The Frequency Distribution of Consolidated Scores

Group of Scores	No. of Students	%
< 40	5	1.53
40～	1	0.30
50～	27	8.26
60～	96	29.36
70～	134	40.98
80～	51	15.60
90～100	13	3.97
Total	**327**	**100.00**

Table 2. The Weights for 16 Indicators

Name	**Meaning**	**Loads** a_{1i}	**Weights** W_i
X_1	Practice1	0.14	0.037
X_2	Practice 2	0.18	0.047
X_3	Practice 3	0.25	0.066
X_4	Practice 4	0.25	0.066
X_5	Practice 5	0.28	0.073
X_6	Practice 6	0.29	0.076
X_7	Practice 7	0.30	0.079
X_8	Practice 8	0.30	0.079
X_9	Practice 9	0.25	0.066
X_{10}	Practice 10	0.28	0.073
X_{11}	Midterm practice exam	0.30	0.079
X_{12}	Final theory exam	0.33	0.087
X_{13}	Final practice exam	0.27	0.071
X_{14}	Attendance	0.16	0.042
X_{15}	Independence	0.20	0.052
X_{16}	excellent homework	0.03	0.008

From the weight of the 16 indicators of comprehensive evaluation in Table 4 we can see that the weights of midterm practice exam, final theory exam and final practice exam are 0.079, 0.087 and 0.071, respectively. They have greater weights compared with other indicators. It indicates that the 3 exams account for a large proportion in the entire comprehensive evaluation. This also matches objectives of the exams. The consolidated scores of 327 students f_1, f_2 ,…, f_{327} can be computed from the weights in Table 4. The frequency distribution of consolidated scores is shown in table 3.

The normal test for the distribution of scores shows: The kurtosis coefficient is -1.243 (standard error =0.135, $P < 0.001$); The skewness coefficient is 6.768 (standard error =0.269, $P < 0.001$). That is the consolidated scores do not fit normal distribution at the test level of 0.05. We checked the values of the 16 indicators for the students whose consolidated scores less than 60 points and found that the 5 students whose consolidated scores less than 40 did not attend almost all lectures and their normal operating performance 0 points. Their low consolidated scores make the whole consolidated scores do not fit normal distribution. If we remove the consolidated scores of these 5 students, then the kurtosis coefficient is -0.077 (standard error =0.271, $P > 0.05$); The skewness coefficient is 0.172 (standard error =0.136, $P > 0.05$). That is the consolidated scores fit normal distribution at the test level of 0.05.

The average scores of 16 indicators for the 33 students whose consolidated scores less than 60 points (failing) and the 13 students whose consolidated scores more than 90 points (excellent) computed using our method are shown in Table 4.

Table 4. The average scores for consolidated scores less than 60 and more than 90

Indicators	< 60		> 90	
	Average	Rate	Average	Rate
X_1	0.44	0.037	4.62	0.385
X_2	0.19	0.016	3.31	0.276
X_3	0.34	0.028	6.14	0.511
X_4	0.25	0.021	5.85	0.487
X_5	0.31	0.026	9.23	0.769
X_6	0.34	0.014	11.58	0.482
X_7	0.47	0.019	12.77	0.532
X_8	0.47	0.019	9.73	0.405
X_9	0.23	0.019	5.54	0.462
X_{10}	0.22	0.018	6.63	0.552
X_{11}	7.97	0.016	321.54	0.642
X_{12}	22.31	0.223	81.27	0.813
X_{13}	1.56	0.010	75.38	0.502
X_{14}	45.31	0.453	49.62	0.496
X_{15}	56.56	0.566	93.08	0.931
X_{16}	96.25	0.962	98.46	0.985

From the results of Table 4 we can see that for the students classified as failing, their rates of the 10 practice works and mid-term and final exams do not account for 0.05, reflecting that the actual levels of these students were low. For the students classified as excellent, their rates of the 10 practice works and mid-term and final exams approximated to 0.5, reflecting that the actual levels of these students were better. Therefore, the comprehensive evaluation of the course determined by our method is in line with the actual situation.

4. Concluding Remarks

The content of above discussions is one of a series of achievements of teaching reform for the national excellent course "Algorithms and Data Structures". The most valuable and most difficult reform is to strengthen the practical aspects of the curriculum, while strengthening the teacher-student interactive teaching. Another innovation closely related to this teaching reform is a comprehensive scientific evaluation of teaching methods for this complex teaching process. This evaluation method will evaluate the whole process of teaching by a dynamic multi-factor evaluation. It will effectively promote the teaching level continues to increase.

Our course group has programmed our comprehensive evaluation method into computer software with C++ language. The software is easy to use and interface friendly. You can easily calculate each student's comprehensive scores after inputting the 16 values of indicators of all students. To make the software more flexibility and adaptability we added some parameters which can be chosen according to actual situation, such as lowest score and highest score on consolidated scores. The rate of failure could also be set in advance in accordance with the teaching requirements. Our course group has also compared the results computed by our software with the results computed by the world's leading software SAS (Statistical Analysis System) [3, 4]. The results are consistent. This proves the correctness of our method and our software from one side.

References

[1] ACM/IEEE-Curriculum 2005 Task Force. Computing Curricula 2005, Computer Science. IEEE Computer Society Press and ACM Press, September 2005. (http://www.acm.org/education/curric_vols/CC2005-March06Final.pdf)

[2] H. W. Marsh, Stability of individual differences in multi-wave panel studies: Comparison of simplex models and one-factor models. Journal of Educational Measurement, vol. 30, pp. 157-183,1993.

[3] R. P. Cody and J. K. Smith, Applied Statistics and the SAS Programming Language (5th Edition), Prentice Hall, 2005, pp. 386-393.

[4] R.A. Johnson, and D.W. Wichern, Applied multivariate statistical analysis. 4thed. Beijing: Tsinghua University Press, 2001, pp. 327-328.