Development of the Multi-Criteria Evaluation Model

Aleksandar VUJOVIĆ¹, Rade GRUJIČIĆ¹, Branko ANĐIĆ², Zdravko KRIVOKAPIĆ¹, Snežana GRUJIČIĆ³ and Jelena JOVANOVIĆ¹,*

¹ University of Montenegro, Faculty of Mechanical Engineering, Podgorica, Montenegro
² University of Montenegro, Faculty of Natural Sciences and Mathematics, Podgorica, Montenegro
³ Secondary Vocational School Pljevlja, Pljevlja, Montenegro

*Corresponding author

Keywords: Customer satisfaction, Evaluation model, Knowledge, Multi-criteria evaluation, Quality.

Abstract. In order to achieve a higher competitiveness and quality level, development of various models including multi-criteria evaluation models is of particular importance for the organization, regardless of activity and size. The knowledge development, as particularly significant variable for the overall quality and performance of the organization, more recently represents a preoccupation of many studies. This paper is done in that direction, as an authors’ tendency to provide an adequate model for the knowledge level evaluation. Some of the major evaluation problems are untimely and unsubstantiated performance rating, a small number of evaluating criteria, the same rating of the criteria with different importance. The authors proposed an evaluation model, aimed at alleviating or eliminating the above-mentioned problems, and then applied it in the performance assessment for secondary school students. The approach is based on a multi-criteria assessment and continuous monitoring, recording and presentation of students’ activities and results through the specially prepared database. The evaluation objectivity and transparency and students’ satisfaction with the evaluation process are increased in this way. Furthermore, students’ work accountability was encouraged, in order to prepare them for the graduation exam and following educational levels.

Introduction

Acquiring a quality education at university and pre-university level is directly correlated with further success in the labor market and strengthening the overall competence of organizations that employ graduates and high school students. Many studies on this subject indicate the importance of ensuring customer satisfaction in the education system, precisely for the reason of creating a competitive business environment [1]. Customer satisfaction with the education system, as well as the improvement of education system quality, affect many other factors, such as the improvement of psychological and sociological conditions, as a precondition for strengthening the overall economic environment [2-4]. These are the aspects which point to the need of making this paper, which seeks to develop a unique model for quality performance improvement in order to ensure customer satisfaction and create preconditions for improving the business environment and competitiveness.

The evaluation procedure is one of the most complex educational activities used to determine the level of knowledge or outcomes achievement in the learning process. The evaluation system is what affect the teaching process and knowledge quality improvement [5]. Many studies indicate that the biggest stress among students, as the service users, is caused by the evaluation procedure [6,7]. In fact, a constant fear of making mistakes and checking the achievement level lead the educational services users to make mistakes, or resort to improvisation [8]. Therefore, it is necessary to pre-define the conditions and make a recognizable evaluation model which is acceptable for users [9]. Many deficiencies in the existing evaluation models can be seen through: evaluation of discontinuities and unrelated learning, subjectivity, lack of clearly defined criteria, unclearly defined learning process outcomes and measures of knowledge level [10,11]. These and other mentioned reasons represent the main motive for creating the multi-criteria evaluation model, which is presented in this paper.
Experimental Research and Methodology

The method of theoretical analysis, descriptive method, didactic experiment method, survey techniques and statistical analysis were used for preparation of this paper. Up to now published papers on the subject of evaluation and docimology were considered by the theoretical analysis method, after which the database with criteria for evaluating students was created. Each of the criteria carries a certain weighting factor that complies with the graduation exam, literature, learning outcomes and Bloom’s taxonomy. In such defined evaluation framework, student gets a mark for each criterion, wherein the final students’ mark for evaluation period represents the nearest integer value of an auxiliary value $M$, which is defined as the sum of the multiplication of marks for each criteria with appropriate weighting factors divided by the sum of all weighting factors.

Based on the previously described performance evaluation procedure, schematic view of the entire evaluation process is presented in Figure 1.

For the purposes of this research, a database was created, so teacher continuously enter the following information in database as evaluation criteria: individual work on class, participation in group work, test, written assignments and essays, students’ behavior and interest (Figure 2). The database was created in Microsoft Excel.

This database was created for the purpose of teaching the subject Montenegrin-Serbian, Bosnian and Croatian language and literature. Due to the specifics of the subject, the largest weighting factor 10 was given to the test which covers all course contents. Previous studies also indicated the importance of the tests. It is particularly necessary to use tests for measuring of students’ advanced mental skills [12]. Test is followed by the written assignments and essays with a weighting factor of 9. Writing should be viewed as an art of expressing what we think, hear and see, so students have to develop this type of communication in themselves [13]. Good writing expression is mainly affected by the writer’s purpose and perspective, but also by the subject, language and style characteristics he/she uses [14]. Individual work at any class carries a weighting factor of 8. It includes the students’ interest and activity during the teaching process, commenting read labor, analysis of literary characters..., whereby students are directed toward critical reading and thinking. Team work is valued with weighting factor of 6. Increasing the interest in the work teams’ performance has a main role in changing processes in the last few decades [15]. The final mark also takes into account students’ behavior and interest, but with the lowest weighting factor – only 4.

After creating, the database was applied in practice for the students’ evaluation in the Secondary Vocational school in Pljevlja, within the subject Montenegrin-Serbian, Bosnian and Croatian language and literature. After the second evaluation period in 2014/15 school year, a surveys of students was performed with the aim to review their attitudes on this evaluation model. A total of 200 students were surveyed. The survey was consisted of 18 questions, 12 closed and six open-ended. In
order to monitor the customer satisfaction level, the research was repeated in the 2016/17 school year. The same students participated, 147 of them, and the survey questions were the same.

Results and Discussion

The survey results pertaining to closed questions are presented in Table 1, while Table 2 shows the most common answers to open-ended questions.

Table 1. The survey results pertaining to closed questions.

<table>
<thead>
<tr>
<th>Question</th>
<th>18/19</th>
<th>mostly no</th>
<th>more or less</th>
<th>mostly yes</th>
<th>yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you understand the essence of the evaluation approach to be applied?</td>
<td>14/15</td>
<td>2.00</td>
<td>2.00</td>
<td>4.00</td>
<td>20.50</td>
</tr>
<tr>
<td>Does application of this evaluation model provide a clear picture of</td>
<td>14/15</td>
<td>1.00</td>
<td>2.50</td>
<td>6.50</td>
<td>19.00</td>
</tr>
<tr>
<td>what you are expected as student within a given teaching subject?</td>
<td>16/17</td>
<td>0.68</td>
<td>0.68</td>
<td>3.40</td>
<td>17.69</td>
</tr>
<tr>
<td>Does this evaluation model allow full and timely insight into your</td>
<td>14/15</td>
<td>0.00</td>
<td>3.50</td>
<td>7.50</td>
<td>31.00</td>
</tr>
<tr>
<td>own achievements in all aspects?</td>
<td>16/17</td>
<td>0.68</td>
<td>0.68</td>
<td>3.40</td>
<td>21.09</td>
</tr>
<tr>
<td>Does this evaluation model solves students’ problem embodied in the</td>
<td>14/15</td>
<td>5.00</td>
<td>2.50</td>
<td>11.00</td>
<td>25.50</td>
</tr>
<tr>
<td>sentence: “I know better than another student who has the same mark</td>
<td>16/17</td>
<td>0.68</td>
<td>1.36</td>
<td>5.44</td>
<td>23.13</td>
</tr>
<tr>
<td>as I”?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is different weighting of criteria the advantage of this evaluation</td>
<td>14/15</td>
<td>0.50</td>
<td>2.50</td>
<td>13.00</td>
<td>30.50</td>
</tr>
<tr>
<td>model?</td>
<td>16/17</td>
<td>1.36</td>
<td>0.68</td>
<td>7.48</td>
<td>22.45</td>
</tr>
<tr>
<td>“Test” is validated as the most important criterion</td>
<td>14/15</td>
<td>4.50</td>
<td>6.50</td>
<td>11.50</td>
<td>30.50</td>
</tr>
</tbody>
</table>
and “Behavior and interest” as the least important one. Do you think if it is correct?

Does this evaluation model encourage the well-timed work of students?

<table>
<thead>
<tr>
<th>Year</th>
<th>Test</th>
<th>Written assignments</th>
<th>Behavior and interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>14/15</td>
<td>4.00</td>
<td>3.50</td>
<td>15.50</td>
</tr>
<tr>
<td>16/17</td>
<td>0.68</td>
<td>1.36</td>
<td>6.12</td>
</tr>
</tbody>
</table>

Does this approach to evaluation enable an introduction to quality and dedication of your own work at every single class?

<table>
<thead>
<tr>
<th>Year</th>
<th>Test</th>
<th>Written assignments</th>
<th>Behavior and interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>14/15</td>
<td>4.00</td>
<td>3.50</td>
<td>15.50</td>
</tr>
<tr>
<td>16/17</td>
<td>0.68</td>
<td>1.36</td>
<td>6.12</td>
</tr>
</tbody>
</table>

Does this evaluation model enable better insight into all the factors that lead to the final mark?

<table>
<thead>
<tr>
<th>Year</th>
<th>Test</th>
<th>Written assignments</th>
<th>Behavior and interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>14/15</td>
<td>2.00</td>
<td>2.00</td>
<td>10.50</td>
</tr>
<tr>
<td>16/17</td>
<td>0.68</td>
<td>0.00</td>
<td>6.12</td>
</tr>
</tbody>
</table>

Is the final mark in this subject fully explained by using this approach?

<table>
<thead>
<tr>
<th>Year</th>
<th>Test</th>
<th>Written assignments</th>
<th>Behavior and interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>14/15</td>
<td>2.00</td>
<td>2.00</td>
<td>10.50</td>
</tr>
<tr>
<td>16/17</td>
<td>0.68</td>
<td>0.00</td>
<td>6.12</td>
</tr>
</tbody>
</table>

What is your current mark of this subject?

<table>
<thead>
<tr>
<th>Year</th>
<th>Test</th>
<th>Written assignments</th>
<th>Behavior and interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>14/15</td>
<td>13.00</td>
<td>44.00</td>
<td>19.00</td>
</tr>
</tbody>
</table>

In your own discretion, rank the significance of usually valued categories on a scale of 1 to 5:

- **Test**
  - 2014/15: 5 (59.50%)
  - 2016/17: 5 (35.37%)

- **Written assignments**
  - 2014/15: 3 (31.50%)
  - 2016/17: 3 (31.50%)

- **Behavior and interest**
  - 2014/15: 2 (30.00%)
  - 2016/17: 1 (60.54%)

Table 2. The most common answers to open-ended questions

It is a common phenomenon in the classical evaluation system that students who are not ready for the tests or written assignments leave the classes in order to avoid the failing marks. Is that the case with this subject? Explain your answer.

- **Students do not deliberately leave the classes without reason.**
  - 84.50 [14/15] 42.18 [16/17]

- **In this way students realize that they can get a higher mark at any time and achieve better results when they are present than when they are absent.**
  - 7.50 [14/15] 23.81 [16/17]

Share your opinion about the fact that each evaluating segment is accompanied by teacher’s comments and notes on the basis of which the final mark is derived.

- **By using comments nothing is forgotten while any doubts are removed.**

- **Students become aware of how this model is good when gain insight into comments.**
  - 18.00 [14/15] 17.01 [16/17]

- **By precisely specified criteria, marks become real and fully explained.**
  - 18.00 [14/15] 17.01 [16/17]

- **I know my mark at any moment and the mark I can strive to.**
  - 4.50 [14/15] 12.93 [16/17]

What do you consider as the most valuable characteristic of this evaluation model?

- **The most valuable feature of this evaluation model is the objectivity / fairness / reality in evaluating knowledge.**
  - 40.50 [14/15] 29.25 [16/17]

- **The most valuable characteristic is the fact that all students are equal, while evaluation is transparent and fully explained, so we have no objections at each other.**

- **The most valuable feature is the multi-criteria evaluation in which the criteria are scored differently, and marks are made on the basis of comments.**

- **There are various ways to achieve the desired results and get the higher marks.**
  - 9.50 [14/15] 8.16 [16/17]

What do you see as the main weakness of this kind of evaluation?

- **This evaluation model doesn’t have weaknesses.**

- **No possibility of back-answering and repairing marks from the previous periods.**

- **A little relaxation and idleness may negatively affect the evaluation.**
  - 3.50 [14/15] 8.84 [16/17]

- **There is no favoritism nor donation of marks.**
  - 3.00 [14/15] 4.08 [16/17]

This kind of evaluation model is best suited for which type of students, in your opinion?

- **This evaluation model fits all students.**
  - 35.50 [14/15] 44.90 [16/17]

- **This evaluation model meets students who are active at the classes and work on time.**
  - 35.00 [14/15] 29.25 [16/17]

- **This evaluation model corresponds to weaker and withdrawn students, considering that their least efforts are being valued, enabling them to come to the fore.**

What should be changed in order to make this model better and more acceptable for pupils?

- **Nothing needs to be changed, because everything is correct / quality / objective /**
  - 65.00 [14/15] 52.38 [16/17]
Remedial tests should be organized. This evaluation model should be introduced into education system from the very beginning of schooling, because it would cultivate and appreciate the work habits. The average rating of the first ten closed questions is 4.39 in the survey conducted in 2014/15, and 4.64 in the survey conducted in 2016/17. These ratings can be considered as a students’ attitude quantitative measure in relation to this evaluation model (Figure 3.a). Their comparison with the average students’ marks at the classification periods after which the surveys were conducted, which was only 2.63 in 2014/15, ie. 2.82 in 2016/17 (Figure 3.b), points to the fact that even those students who have a poorer assessments of a given subject experienced this evaluation model as a high quality. They were motivated to work because it was perceived and evaluated from several aspects. This conclusion correlates with previous research. If students are rewarded for the results they achieve in both verbal and writing tasks and encouraged to express creativity in work, assume responsibility and participate in the decision-making process they would be motivated for work [16]. Answers to the question No. 15 also suggest such conclusion.

Those characteristics that some students experienced as negative actually represent the quality of this evaluation model. If student is not accustomed to responsibility and seriousness in the work, he/she will experience the need for the continuous and transparent learning as the downside of this evaluation approach. Hence 3.5% of students gave the following answer on the question No. 16: “A little relaxation and idleness may negatively affect the evaluation”, while 3% of them is not satisfied with exclusion or substantial decrease of teachers’ subjectivity. These percentages are even higher after repeated survey (8.84% and 4.08%), although it was expected that the students will get accustomed to the necessity of regular work and grades distribution by merits. If this evaluation model become a part of the teaching practice in the lower education levels, students’ work habits will undoubtedly be formed. A similar conclusion was carried out by the students. Even 18.37% of them at the repeated survey said that this evaluation model should be used from the very beginning of education in order to properly prepare them for university education.

In order to monitor the level of customer satisfaction, a comparison of survey results for the school year 2014/15 with the results for the school year 2016/17 is made. For each of the survey questions Pearson’s correlation coefficient was calculated. It turned out that there is very high correlation between the results for each of the questions, with the average correlation coefficient value of 0.95.

By comparing the pie charts given in the Figure 3.a, ie. marks on the closed questions given in Table 1, it can be seen that the total number of the maximum marks (5) on closed questions is significantly increased on the repeated survey. In addition, the average score on the closed questions
(previously adopted measure of students’ satisfaction with evaluation model) is increased from 4.39 to 4.64, which indicates that customers’ (students in this case) satisfaction level has been uplifted.

By analyzing the grades that students achieved after the second classification period of 2014/15 academic year and the second classification period of 2016/17 academic year (question No. 11 in Table 1), it can be seen that the number of insufficient marks remained at the same level, the number of sufficient marks was decreased, while the number of good marks was increased and the number of very good marks was decreased, while the number of excellent marks was increased. Such a situation indicates that some students who have made an effort increased their marks, so the overall average subject grade at the level of class rose from 2.63 to 2.82.

Conclusion

Many years of teaching and evaluation practice pointed out that one of the biggest evaluation problems are the untimely and unfounded evaluation, too few criteria that affect the marks, the same treatment of criteria of different significance, completely ignoring or overly emphasizing of the educational process behavioral component. Therefore, multi-criteria approach based on continuous monitoring and recording of student activities was introduced to eliminate or mitigate the above mentioned problems.

Feedback from students indicated that, regardless of the level of their success in the teaching process, objectivity, openness, transparency and the daily record of student work are seen as the values of the multi-criteria evaluation approach, enabling them to continuously monitor their own advancement. They pointed out that the recording of comments on each form of student work allows timely identification and elimination of shortcomings in their own learning, but also helps them to understand learning outcomes through defining weighting factors for all the teaching content and teaching methods. The behavioral component of the educational process has its own place in this evaluation model, encouraging the work responsibility and regularity and, above all, build work habits necessary to achieve good results, both at the future educational and professional level.

References


[5] JISC (2010). Effective Assessment in a Digital Age; A guide to technology-enhanced assessment and feedback. HEFCE


