ASSESSMENT OF QUANTITATIVE AND AUTOMATED RUBRICS AS LEARNING TOOL IN ENGINEERING THERMODYNAMICS

Begoña Peña, Manuel Bailera

Department of Mechanical Engineering, Universidad de Zaragoza (SPAIN)

Abstract

Evaluation is an essential step in the teaching-learning process, both to monitor and improve the process itself and to assess the learning performance of each particular student. The use of evaluation for formative purposes motivates a higher engagement of the student and helps to detect deviations in the acquisition of competences that can be solved in time.

Rubrics are defined to structure the expectations for an assignment by listing the assessment criteria and by describing levels of accomplishment. The criteria are the particular skills to be evaluated; the levels of performance are ratings to distinguish the quality of achievements; descriptors are a specific text for each level and criterion describing what the student must know to do in each case. If criteria, levels and descriptors are known by the student in advance, evaluation by rubrics can be used as a powerful learning tool for self-evaluation and, therefore, could serve to motivate and to promote success.

As occurs for any assessment procedure, rubrics must be defined to be objective, coherent with the competences to be acquired by the student and low time-consuming for the teacher. To design quality rubrics which fulfill these requirements is a very complex task, as they must be specific of the subject, level and context of application. However the biggest challenge is to develop quantitative and automated rubrics to reduce the time invested by the teacher.

In the present work, financed by the Program for Promoting Teaching Innovation (PIIDUZ_18_102), rubrics have been designed for the subject of Engineering Thermodynamics of the Degree in Industrial Technologies Engineering at the University of Zaragoza (Spain). The criteria are the particular skills related to the competences to be acquired in the subject and four levels of performance are defined by specific descriptors. The quantitative information, stored in Moodle forms, is downloaded in spreadsheets and later processed to calculate the score reached by the students in every criterion at different times of the semester. A preliminary assessment about the correlation between rubrics and final success is here presented.

Keywords: Rubrics, Formative Evaluation, Self-assessment, Higher Education Innovation, Engineering Thermodynamics.

1 INTRODUCTION

For the learning success, continuous assessment should be adapted to both subject and students. Students may assess their learning process themselves if provided with clear evaluation criteria, quick feedback on mistakes and knowledge gaps, and suggestions for improvement. In this way, continuous assessment would become a useful learning tool for the acquisition of competences. In addition, it would give teachers valuable information on students’ progress to correct any deviation.

To perform this kind of instructive assessment, rubrics are one of the most suitable approaches. Rubrics are documents that articulate the expectations for an assignment by providing the assessment criteria, and describing the levels of achievement that students may reach while addressing these criteria [1]. Bringing assessment closer to the students helps them to understand the learning objectives, and the competences and the level of performance they need to reach. If students have this information in advance, they can assess themselves to reflect on their current performance and how it could be improved.

Depending on the rubric used, the assessment may be holistic or analytical [2]. Holistic rubrics are easier and quicker to design and give the teacher a global picture of the learning. However, they provide less specific feedback to students, what limits its use as learning tool or self-assessment. Analytical rubrics arrange the assessment in explicit criteria, providing detailed information about the
learning process and recommendations for improvement. The design and implementation processes are more time-consuming for the instructor, but the results are more valuable for the students.

To keep the time spent within reasonable limits when assessing large groups of students, quantitative and automated rubrics are desired. In technological, technical and health areas, quantitative or short-answer questions are commonly used in the evaluation procedure and can be implemented in Virtual Learning Environments (VLE) to grade automatically. If these records are then linked to specific criteria of the rubric, the evaluation is automated.

In this work, the methodology for designing quantitative and automated rubrics is presented. Quantitative results come from Moodle’s assessment activities, while the algorithm was implemented through MATLAB® software. Results include the rubrics developed to give detailed information on what is expected to be learnt in the subject of Engineering Thermodynamics. Moreover, an analysis on the consistency between the learning activities and the competences, derived from the feedback of the automated process, is presented. This study has been carried out in the framework of the teaching innovation project “Combining the flipped classroom model with the use of ICT and active methodologies”.

2 METHODOLOGY
This section presents guidelines for the design of quantitative rubrics, and the automation of the assessment process through a dedicated algorithm.

2.1 Design of the quantitative rubric
A wide variety of tools for developing rubrics are available on the Internet (e.g., pre-designed rubric banks, rubric templates and rubric generators) [2], and many examples of application at every educational level can be found in literature [1], [3]. Moodle [4], among others VLE, also include rubric editor through the Advanced Grading tool. However, the assessment of large groups of students is not affordable by using these methodologies.

From the theoretical point of view, the design of a rubric must include (i) the assessment criteria, (ii) the levels of attainment, and (iii) the descriptors [2]. The assessment criteria are the competences on which students are evaluated; the levels are the different labels that rate the performance of the student (qualitative or quantitative); and the descriptors are short texts explaining what students need to know for each criterion and level.

The correct definition of these elements (criteria, levels and descriptors) is essential for the success of the formative aim of the rubric [5]–[7]. To this end, Dornish and McLoughlin [2] summarised common flaws to avoid while designing rubrics. The most relevant is the utilisation of criteria that are too general, too numerous, or not significant regarding its relative weight in the subject. Levels of quality with unclear differentiations, as well as descriptors too general or too specific, should also be avoided. Panadero and Jonsson analysed 21 studies on the formative approach of rubrics [3]. They concluded that higher transparency is needed in the assessment to make the student fully aware about what is expected from her/him. Also, reducing the anxiety about mandatory tasks, aiding the feedback process, improving student’s self-efficacy and supporting student’s self-regulation, leads to an improvement in the student’s performance and her/him learning process.

Based on these studies, the procedure followed for the design of the quantitative rubrics has been the following:

1. To select the competencies that the student must develop in the subject according to the curriculum.
2. To define the criteria associated with the selected competences. If necessary, several criteria are defined for a competence. It may also be the case that one criterion refers to several competences.
3. To establish four levels of performance: Beginner, Basic, Intermediate and Advanced.
4. To define a specific descriptor for each level of each criterion describing in detail what the student needs to know. The knowledge required to reach subsequent levels is cumulative, so the descriptors include the requirements of previous levels.
2.2 Automation of the assessment by rubrics

The assessment process by rubrics has been structured in the following way:

1. To select the learning activities whose scores will be used to compute the grade in a particular rubric and moment of the academic course.
2. To relate each question to one or more criteria, and to the level of performance within those criteria.
3. To allocate a specific weight to each question within the corresponding criterion and level.
4. To develop an algorithm that computes the student’s grade on each criterion, accounting the scores of the related questions.
5. To download from the VLE the scores of each learning activity in spreadsheets and apply the rubric algorithm.
6. To use the assessment results to periodically provide the rubrics’ outcome to the students. Results may be updated at different milestones during the course (e.g., before an exam).
7. To implement corrective measures in the learning process if a general trend of low scores is observed in a criterion.

The first four steps should be performed before the academic course begins, while steps 5 to 7 are carried out every time the instructor wants to provide updated outcome to the students. The algorithm was programmed in MATLAB® because it has a great flexibility for implementing modifications in case of the decisions taken in steps 1 to 3 are changed during the course or in subsequent courses. The algorithm’s flowchart is included in Fig. 1.

The scores from the learning activities and the list of students are downloaded in spreadsheets from Moodle. These data are previously revised and taken as inputs for the algorithm. In the first part of the code, the user establishes the scope of the rubric according to the stage of the course, determining the criteria that can be assessed and the data to be used. Then, the spreadsheets containing the scores of each activity are read and stored in independent array type variables. Afterwards, some array variables related to intermediate calculations and the structure array variable in which grade are gathered for each student are sized and initialized.

In the second block of the code, the score in the different levels of performance is computed for each criterion as the weighted average of the scores in the corresponding questions. Matlab’s ability to work effectively with arrays has been used instead to perform the mathematical operations with a loop for. To reach a certain level, the qualification must overcome a threshold. If a level of performance is not passed (i.e., the “No” option of the rhombuses in Fig. 1), the algorithm stops and the performance is set equal to the last level successfully achieved. The outcome of each criterion is gathered in the rubric structure array for every student.

In the final step of the process, this information is used for different purposes: feedback for the student about her/his learning progress (even regarding the rest of students), identification by the instructor of learning difficulties in specific criteria and reflection about the assessment process itself to improve the activities, modify the weights of questions and adjust the thresholds to overcome each level of performance.

3 RESULTS

The methodology described above has been applied to the course on ‘Engineering thermodynamics and fundamental of heat transmission’, within the 3rd semester (2nd year) of the B.Eng. in Industrial Technologies Engineering at the University of Zaragoza. This course has 6 ECTS, equivalent to 150 hours of work (45 h of theory, 15 h of laboratory and 90 h of personal study). In the next subsections, the rubric developed for the course and the feedback coming from its automation are presented.
3.1 Quantitative rubric for Engineering Thermodynamics

The competence to develop within the course and the expected learning results are included in the Degree’s curriculum. The competence has a general approach and specifies the following: “Knowledge on applied thermodynamics and heat transfer. Basic principles and their application to the resolution of engineering problems”. The expected learning results are more definite and they are divided in five items: (i) the student knows the thermophysical properties that are relevant in the industrial sector, and is able to identify and use the proper models and tools for their calculation, (ii) the student knows and applies correctly the laws of thermodynamics for the energy analysis of equipment and basic process of engineering, (iii) The student knows the fundamental principles for the analysis of thermodynamic cycles, (iv) the student knows to apply basic heat transfer models for the analysis of thermal equipment, and (v) the student solves basic problems of technical thermodynamics and heat transfer in a reasoned way in the field of engineering.
Based on these requirements, 10 criteria have been defined with 3 learning levels each (if the student does not pass the Basic level, it is assumed a “Beginner” level of performance). The quantitative threshold to pass each level will be adapted to be motivating for the students, according to the learning activities of the course. Table 1 shows various examples of the criteria constituting the rubric designed for the current course on ‘Engineering thermodynamics and fundamentals of heat transmission’.

Table 1. Some example of criteria, levels of performance and descriptors of the rubric designed.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Basic</th>
<th>Intermediate</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manage of properties and processes</td>
<td>The different types of properties and processes are known.</td>
<td>The type of a process is identified correctly.</td>
<td>The equations are applied correctly and the processes are depicted graphically.</td>
</tr>
<tr>
<td>Correct use of Ideal Gas model</td>
<td>The equation of state is applied correctly (units).</td>
<td>The variations of internal energy, enthalpy and entropy are computed.</td>
<td>The Ideal Gas model is applied only under the proper hypotheses.</td>
</tr>
<tr>
<td>Correct use of energy and entropy balances in steady conditions</td>
<td>How to apply the energy and mass balances is known.</td>
<td>How to apply the entropy balance is known.</td>
<td>The balances are applied only under the proper hypotheses.</td>
</tr>
<tr>
<td>Basic knowledge on thermodynamic cycles</td>
<td>Thermodynamic analysis is applied to cycles and their performance is calculated.</td>
<td>The equipment required in a cycle and their roles are known.</td>
<td>The thermodynamic cycles are properly identified and depicted graphically.</td>
</tr>
</tbody>
</table>

3.2 Feedback from automated rubric: consistency between competences and learning activities

The scores of the following learning activities have been used in the automated rubric: 3 questionnaires after the visualization of learning videos (13 questions), 8 questionnaires for self-assessment (102 questions), 6 response forms related to mentored tasks (85 questions) and 5 questionnaires on laboratory work (30 questions). In total, there are 230 questions. The number of questions associated with each criterion and level is shown in Table 2. As some of the questions require applying several concepts to answer, their score is used in more than one criterion. That is the reason why the total number of questions in Table 2 is greater than the actual number of questions that compose the learning activities.

From the allocation of the questions among criteria and levels, it can be observed that the learning activities are not balanced with the criteria that students must reach. As some criteria have significantly fewer questions assigned to them, the maturity that students can achieve in the corresponding knowledge is limited (e.g., in criteria 4, 7 and 9). This feedback coming from the automated rubric design is valuable information for the instructor to improve the learning process. In the next course, new questions related to these criteria will be added to the learning activities.

In some other cases the number of questions is totally justified. In criteria 1, 2 and 5 there are a high number of questions because they are associated to essential concepts for the subject. In criterion 10, the amount of questions is very limited, as they are in fact long problems requiring several mathematical steps. The solution for the latter case is to divide each problem in several short questions or to assign the problems a differentiated weight within the criterion.
Table 2. Number of questions associated with each criterion and level of performance.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Basic</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>35</td>
<td>32</td>
<td>56</td>
<td>123</td>
</tr>
<tr>
<td>02</td>
<td>25</td>
<td>15</td>
<td>22</td>
<td>62</td>
</tr>
<tr>
<td>03</td>
<td>13</td>
<td>6</td>
<td>7</td>
<td>26</td>
</tr>
<tr>
<td>04</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>05</td>
<td>27</td>
<td>11</td>
<td>18</td>
<td>56</td>
</tr>
<tr>
<td>06</td>
<td>9</td>
<td>10</td>
<td>20</td>
<td>39</td>
</tr>
<tr>
<td>07</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>19</td>
</tr>
<tr>
<td>08</td>
<td>26</td>
<td>21</td>
<td>13</td>
<td>60</td>
</tr>
<tr>
<td>09</td>
<td>4</td>
<td>14</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>162</td>
<td>122</td>
<td>153</td>
<td>437</td>
</tr>
</tbody>
</table>

In the present study, all the criteria have the same specific weight. However, they do not have the same relevance regarding the final grade of the course. For the next academic year, this topic will be discussed with the rest of the professors to establish differentiated weights.

4 CONCLUSIONS

In this work we have presented the methodology to define quantitative rubrics and the automation of their assessment through a dedicated algorithm by using the answers gathered from Moodle activities. The methodology has been applied to the course on ‘Engineering thermodynamics and fundamentals of heat transfer’, within the B.Eng. in Industrial Technologies Engineering at the University of Zaragoza. So far, 10 criteria have been defined with their corresponding levels and descriptors, and the algorithm that automatically calculates the score has been developed. Although, during the course it was not possible to inform about the rubric grading to the students because of time limitations, they had the rubric.

From a preliminary analysis of the raw data, it is concluded that some habits of the students could affect the effectiveness of rubrics to predict final success. From a qualitative review of the questionnaires, it has been observed that some students (mostly those repeating the course) barely participate in this type of activity, but then they pass the final test successfully. Moreover, some students fail the final test despite of the good grades obtained in the rubric, as they repeat the questionnaires several times until obtaining the maximum grade (attempts are not limited). This trial-and-error answering leads to deceptive results, since the students have not truly understood the subject. In this regard, it is convenient to inform students about the limitations of the assessment by this kind of rubrics, to make them aware of how they are actually progressing.

Regarding learning outcomes, a higher success rate in the final assessment test (around 15 percentage points) has been obtained with respect to academic year 2017-2018. It is worth to mention that the assessment by rubrics in this study was simultaneously applied with flipped classroom experiences [8] and with a parallel study on learning analytics based on the information gathered in Moodle [9]. For that reason, it is not possible to know to what extent the rubrics have contributed to the positive results. In order to know the opinion of the students in this respect, a survey will be prepared for the next academic year.

Finally, the design procedure of the rubrics described in the present paper had an unexpected usefulness. The association of questions to criteria could reveal lack or excess in time dedication to each criterion. Such previous analysis serves to balance the amount of learning activities associated to each skill or competence. In this study, e.g., it was found that criteria 4, 7 and 9 had significantly fewer questions to study with. Therefore, the number of questions will be increased in the next academic year.
ACKNOWLEDGEMENTS

The authors thank professors Mª Belén Zalba Nonay, Ignacio Zabalza Bribián, Eva Llera Sastresa, and Sergio Usón Gil for their willingness to discuss and share their experience in the field of teaching innovation. This research has been funded by the Program for Promoting Teaching Innovation (PIIDUZ_18_102) from the University of Zaragoza.

REFERENCES


