Case Study Based Design of Prototype Assessment Application

Neville Palmer

School of Media Arts and Technology, Southampton Solent University,
East Park Terrace,
Southampton
neville.palmer@solent.ac.uk

Abstract

When designing assessment mechanisms in higher education a number of challenges should be considered. Assessments must be designed carefully so that students understand what is expected of them and the criteria that will be used to mark them. They also expect timely and meaningful feedback both before and after the assessment cycle that they can use to inform their learning and gauge their level of achievement. Assessment criteria should be designed to be less subjective so that marking can be seen by students to be fair and so that a rationale for marks awarded is clear. If there is little evidence to demonstrate why a particular mark was awarded students may not trust the assessment mechanism. The issues affecting assessment design were investigated with the aim of developing a prototype application that could be used to aid in the design process and also to allow the assessments to be marked. Importantly the prototype is also able to generate a feedback report that clearly informs students of their level of achievement and provides them with both specific and generic feedback that they can use to inform their learning. The prototype was tested on the final year project unit of a computer networking degree programme. This report discusses the important challenges in designing assessments, how the prototype was informed by the issues and how it was tested within the unit at the centre of the case study.

Keywords: assessment criteria, marking, application

1.0 Introduction

Assessment mechanisms for undergraduate students are normally specified when units or modules on a particular degree programme are designed. The specification of Learning Outcomes are of high importance in the design process and the chosen assessment methods must be capable of testing whether students have achieved those learning outcomes. Therefore when the specific assessment instruments are

designed the learning outcomes play an important part in design considerations. Many methods of assessment available and if we are designing assessments for final year undergraduate level students some of these mechanisms are more appropriate than others [1].

When designing assessments one should consider whether the students will understand the assessment criteria. Student diversity may mean that some students understand the assessment criteria better than others. They may have varying educational backgrounds, variations in learning ability and there may be language barriers. Students may have different learning styles and their profile may favour one form of assessment over another [2]. There may be temptations to word assessment criteria in vague and highly academic terms, but there are dangers in this when these factors are considered. At higher levels of study the requirements of assessment might still be quite general, but assessment criteria should still be clearly stated.

Student perspectives on assessment are apparent from surveys, such as the National Student Survey and other internal unit evaluations. Their concerns relate to timeliness of feedback, fairness and accuracy of assessment. If assessment criteria isn't stated clearly there is likely to be more subjectivity in the assessment process and some students may believe that personal factors are taken into account above adherence to criteria. Students prefer written feedback which can help to explain the grade achieved. Students may not always have a full understanding of what constitutes feedback. Some feedback may just be in the form of a tick in a box defining criteria to meet a particular grade even if there are explanatory comments in the box.

Students can participate in the assessment design process. If students have a part in this then they may better understand the assessment, what is required of them and better benefit from self-evaluation [3]. Most engaged students will see the benefit of carefully designed assessment and this may alleviate the concerns of some.

However the assessment criteria should still be carefully constructed. If there is too much subjectivity in the design of assessment criteria it may be difficult for academics to mark it, although some might argue that they can draw on personal experience or academic judgement.

Particularly in the case of final year projects team marking is often involved so it is important to have clear criteria for assessors so that the whole team is marking to a common standard. If the criteria are too vague then the assessors may place their own interpretation on them and this may give rise to wide discrepancies between first and second marker. Students may not have confidence in the assessment process if this is the case and there may be issues for external moderation and quality assurance. Some people may argue that less detail in the definition of assessment may provide scope for students to apply their own academic judgement in the final year of study, whereas too much detail may be prescriptive. However

this may be the case for the description of assessment requirements, but the marking criteria should still be clearly defined. If the criteria are not well defined then students may be unsure of what is expected of them and in turn the assessor may also be unsure. This may introduce more subjectivity into the marking processes and increase the challenges of explaining the grade awarded to a student. If criteria is more detailed students will know more precisely what is expected of them and if the assessor strictly applies the criteria the assessment is more likely to be seen by students as fair and equal [4].

Another factor that may have an impact on clarity of criteria based assessment is whether we define a single or multiple criteria. In some cases, unless carefully designed, a single criteria may be more coarse grained and more subjective than multiple criteria. Sometimes when a single criteria is used it may in fact combine more than one element, in which case it may be better to split the elements into multiple criteria. For example supposing we have a single criteria that is described as "Report that analyses the requirements and presents and justifies solutions for the case study" and the standard to reach first class honours level is described as "Excellent analysis of requirements of the case study. Excellent analysis and justification of solutions that clearly match all requirements of the case study". If a fail grade is described as "No adequate description of requirements. Unable to adequately present technologies" what mark do we award a student who has provided no analysis of requirements at all, but has an excellent justifications of solutions that match the given requirements of the case study? This would be much clearer if we split this up into at least two separate criteria, one for marking the analysis of requirements and the other the justification of solutions.

Another method that might be adopted is to split criteria up further into sub-criteria. For example one criteria in a project report might cover planning and this might be split further into sub-criteria such as planning methodology, management and control, etc.

We need to consider what artefacts the student is submitting as part of the assessment. It may just be a written report or answers to examination questions. However even reports may require accompanying evidence, for example if the report is based around a project or case study then accompanying evidence such as notes, background work or minutes of meetings may be among items included in an appendix and discussed within the main body of the report. A portfolio may consist of evidence of work undertaken with accompanying reflection or discussion. In some types of assessment we may wish to keep track of which items have been included in the final piece of work. This could also take the form of skills based assessments where the assessor may wish to simply tick off items successfully completed. In a student centred learning approach personal learning records may include reflection on learning with supporting evidence. These may cover a number of preparation and laboratory sessions. So we may wish to have one set of criteria, but which can be applied to a number of separate sessions.

Academics should also consider their personal time constraints when designing an assessment in terms of time to develop the assessment, time to run or set the assessment and time to actually mark it. Some assessments may seem the perfect assessment mechanism, but in reality they may not be scalable if large numbers of students are involved. As discussed earlier students expect timely feedback and the detail and timeliness of the feedback that academics are able to provide is a further challenge on their time. Meaningful feedback is time consuming to produce. However if much of the feedback is common among a range of students then it may be appropriate to reuse this. What if the generation of feedback were automated in some way?

2.0 The Application Background

The design of the assessment of projects on the final year of a computer networking degree programme was investigated. In the case of the final year project unit all dissertations are double marked. The course team noticed that there could be fairly wide variations in marks awarded to students by first and second marker. They determined that this was most likely due issues in the clarity of the marking criteria, which meant that academics often had to place their own interpretation on it. This may have been due in part to the fact that a range of different programmes from engineering to computing had originally used the same criteria. When the computer networking course team were free to design their own criteria they were able to make improvements in the clarity. The latest incarnation of the new design for assessment of dissertations involves four criteria, each split into four sub-criteria. This development succeeded in reducing the variation between markers.

A spreadsheet was developed by a member of the course team that enabled team members to mark reports and dissertations for the project using the new criteria. Spreadsheet formula enabled an overall mark to be calculated and it was able to generate generic feedback for students on another page and comments could also be entered manually by the marker.

2.1 The Software

Following experience learned from the development of an application for laboratory automation an investigation was conducted to see whether it was possible to develop an application to automate the process of assessment criteria design and marking, not just for the project unit, but for others too, though the project unit would be used for testing [5]. The laboratory automation application, developed in Visual C#, took the output of information in text and list boxes from a windows form to customise PowerShell scripts to manage a networking laboratory [6]. Perhaps it would be possible to apply a similar approach to customising feedback reports for students from assessment. The aims of the prototype would be to provide a generic assessment design tool for academics that allows detailed assessment criteria to be easily defined, allows student work to be assessed accurately and is able to generate detailed and timely feedback reports for students

to assist in their learning and understanding of achievement. The application would enable common elements of feedback to be re-used.

The key features of criteria based assessment design that must be incorporated into the application are:

1. Grading scheme
2. Learning outcomes for the unit
3. Description of assessment requirements
4. Description of assessment criteria, with criteria for each grade in our grading scheme (this should match the expectations of the learning outcomes).

2.1.1 Grading Schema

In order to mark an assessment a grading scheme is necessary, where we need to allocate a mark according to the value that we place on the work submitted for each criteria [7]. Using percentages as the grading scheme is possible, but if the grade boundaries are in 1% increments this will be too fine grained. It would be difficult to justify why one had awarded one student 67% for a particular criteria and another student 68%. We could employ a scheme whose boundaries increased in 10% increments. However higher education tends to classify degrees as first, upper second, lower second, third and unclassified. This only allows 5 categories in the grading scheme, which is too coarse grained. For example the first class category covers 100 to 70%. So many institutions subdivide the classifications. One method might use A+, A, A- to cover the range between 70 to 100% and so on. The system adopted at the author's university is called Grade Marking and uses A1 to A4 to cover the first class range, B1 to B3 to cover upper second, etc. The boundaries between grade categories are then neither too fine or too coarse grained.

2.1.2 Learning Outcomes

When a particular course is validated each unit is designed around a series of learning outcomes. There may be general learning outcomes for each course, but more specific learning outcomes for each unit. The learning outcomes define the skills that students studying each unit should acquire. Student success in each unit is then gauged according to the skills that they have acquired, so assessment should measure whether a student has achieved the learning outcomes and to what level of competency. Examples of learning outcomes might be Cognitive, Knowledge and Understanding, Professional and Practical skills, etc. For example a cognitive skill might be "to evaluate solutions for a case study", and the case study might be specific to computer networks in this case.

2.1.3 Description of Assessment Requirements

This should provide information on what students are expected to do for the assessment. For example this might require them to investigate a case study and

produce a report on solutions for the case study. This might be more prescriptive in nature for students at lower levels of study, but less so towards post-graduate level. Nevertheless it is important to make it clear to students what is expected of them. Student diversity may require a clear and guided explanation of what is required.

2.1.4 Assessment Criteria

The assessment criteria describes the marking scheme. These criteria should clearly describe what standard of work is expected from students for each grade within the grading scheme. If we employ a single criteria to describe the whole assessment outcomes then this may prove too subjective [8]. Multiple criteria might be employed to enable more fine grained marking. Another technique might be to split the criteria further into sub-criteria.

Once we have designed our criteria for the assessment when marking we could just select the appropriate grade that we think the student has achieved for each of the criteria, the equivalent of ticking a box on a paper form, but generic feedback at least should be provided to assist students with development of their learning. However they value specific comments very highly, but these can be time consuming to produce, whilst students expect a mark and feedback quickly.

2.2 Requirements of Assessment Application

The requirements of the new software needed to take into account the elements of assessment design and the issues discussed above can be summarised as follows:

The application must incorporate:

1. Learning Outcomes to be entered and saved, or loaded for later use from file.

1.1 Each learning outcome to have a title and description.

2. A Grading schema to entered and saved or loaded for later use from file.

2.1 Each grade within the schema to have:

2.1.1 Title,

2.1.2 Generic feedback descriptor,

2.1.3 Percentage equivalent,

2.1.4 Lower percentage boundary,

2.1.5 Upper percentage boundary.

3. An assessment to be defined and saved or loaded from file.

For each assessment the application must incorporate:

3.1 In the description:

3.1.1 Title and code,

3.1.2 Detailed description of the assessment requirements,

3.1.3 Weighting within the unit.

3.2 Learning outcomes must be selected for the whole assessment and optionally for individual criteria within the assessment.

3.3 Multiple assessment criteria and/or sub-criteria to be defined, with:

3.3.1 assessment requirements for each grade within the selected grading schema,

3.3.2 and optional feedback for improvement for each grade.

3.4 Allowance must be made for sub-criteria within each major criteria item.

4. The application must allow marking of students according to the criteria:

4.1	The marking of multiple criteria and sub-criteria by selecting an appropriate grade from the grading schema.
4.2	The optional capability of checking off of each criteria to indicate whether that criteria has been completed.
4.3	The capability of marking of multiple sessions.
4.4	The entry of optional comments by the assessor:
4.4.1	adjacent o each criteria,
4.4.2	and for the overall assessment.
4.5	An overall grade must be generated for the student for the assessment and also an equivalent percentage.
4.6	Student assessment must be saved to file or loaded for later use from file.
4.7	A feedback report in rich text format must be generated and:
4.7.1	saved to file or loaded from file,
4.7.2	with the capability of being printed.
4.8	The feedback report must include:
4.8.1	Student name, assessment title and weight,
4.8.2	Detailed description of the assessment requirements,
4.8.3	Optional learning outcomes for the assessment,
4.8.4	Overall mark from the grading schema,
4.8.5	Optional percentage equivalent,
4.8.6	Overall comments by the assessor on the whole assessment,
4.8.7	For each criteria and sub-criteria:
4.8.7.1	Title,
4.8.7.2	Description of criteria or sub-criteria,
4.8.7.3	Grade awarded,
4.8.7.4	Feedback for the grade,
4.8.7.5	Optional generic feedback for the grade,
4.8.7.6	Optional feedback for improvement,
4.8.7.7	Comments from the assessor on this criteria/sub-criteria.

3.0 Case Study Unit: Project

The prototype was developed to incorporate the key features and requirements. A tabbed layout is used to represent each of the key features: Grades, Learning Outcomes, Assessment. It was tested by configuring it with the requirements of the case study.

3.1 Grading Schema and Grades Tab

The University uses a standard grading schema, Grade Marking and has a generic statement of standards for each grade as follows:

Grade	Equiv%	Generic comments
A1	100	Exceptional in most / all aspects, substantially exceeds expectations for this level
A2	92	“””””
A3	83	Excellent quality, exceeding expectations for this level in many aspects
A4	74	“””””
B1	68	Meets all intended learning outcomes & exceeds threshold expectations for several of them
B2	65	“””””
B3	62	“””””

C1	58	Meets all intended learning outcomes & exceeds threshold expectations for some of them
C2	55	””””
C3	52	””””
D1	48	Meets all required learning outcomes at, but rarely exceeds the threshold expectations
D2	45	””””
D3	42	””””
F1	35	Fails to meet all of the intended learning outcomes and is marginally inadequate for this level
F2	20	Fails to meet all of the intended learning outcomes and is inadequate for this level
F3	15	Fails to meet all of the intended learning outcomes and is inadequate for this level
S	1	Submitted
N	0	Non-submission

The Grades tab was configured with a treeview to represent grades within the grading scheme. Each grade has a title and the generic feedback comments can also be entered. The equivalent percentage must be entered along with the lower and upper grade boundaries for each. The grades are presented on the Assessment form in a selectable listbox. The grading scheme can be saved or loaded from file. Fig. 1 demonstrates how the University Grade Marking scheme was represented within the application.

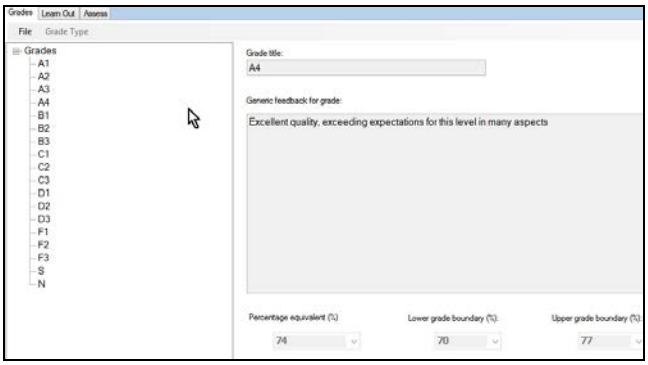

Figure 1: Grading Scheme representation in the application.

3.2 Learning Outcomes and Learning Outcomes Tab

The dissertation element of the final year project has four learning outcomes:

1. Cognitive Skills
    a. Select, apply and evaluate the appropriateness of methods, tools and technologies in the synthesis of meaningful project outcomes.
2. Practical and Professional Skills
    a. Undertake a significant self-managed project in a planned and systematic fashion.
    b. Identify, interpret and integrate theory drawn from a range of appropriate sources.
3. Transferable and Key Skills
    a. Communicate clearly and concisely verbally and in writing

These were configured from the Learning Outcomes tab in the application. These are reflected in a listbox on the Assessment tab so that one or more of the learning outcomes can be selected. The learning outcomes can be saved or loaded from file.

Figure 2 shows how the learning outcomes for the project report were configured within the application.

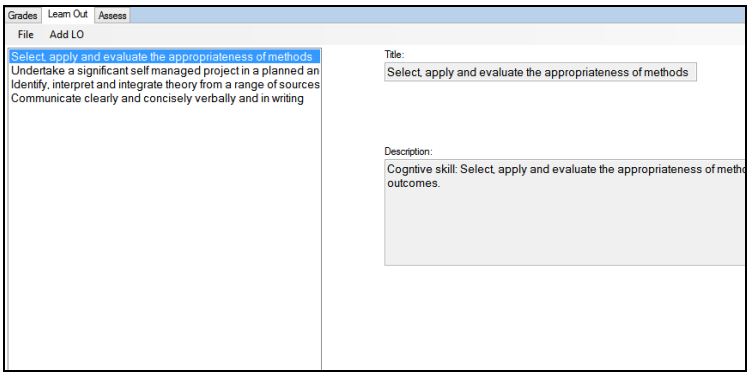

Figure 2: Learning Outcomes tab

3.3 Assessment Criteria and Assessment Tab

The project dissertation assessment designed by the course team included four main criteria, each with four sub-criteria as follows:

1. Product/Results (40%)
    a. Development of consistent & complete project aims, objectives and specification
    b. Selection & application of appropriate tools, technologies & processes
    c. Discussion of issues involved in the design, implementation & testing
    d. Evidences project complexity, completeness & coherence, also reflected in aims
2. Professional Practice (project methodology and planning) (20%)
    a. Discussion & justification of appropriate methodology
    b. Formulated consistent project planning, monitoring & control strategy
    c. Discussion & development of appropriate metrics for project testing
    d. Reflection on development of criteria for success
3. Evaluation (critical review of the project) (20%)
    a. Evaluation of product results
    b. Evaluation of project planning
    c. Evaluation of success in meeting project objectives & specification
    d. Formulate meaningful recommendations
4. Reporting (20%)
    a. Presentation of clear & well structured reporting in which all aspects of project are covered
    b. Appropriate academic style
    c. Integrated a good quality list of references correctly cited
    d. Comprehensively and coherently articulated referenced report

The criteria were weighted according to the percentage shown in brackets, however the sub-criteria within each of the main criteria were evenly weighted (ie. 25% within each criteria).
These criteria were entered from within the Assessment tab, which contains the most complicated part of the application. There are two modes, one an editing

mode in which the assessment is designed and edited and the other a marking mode that allows the assessor to mark a student's work.

When designing the assessment in editing mode a title and weighting within the unit must be included. A description of the assessment must describe what the student is required to do for the assessment and may refer to other documents if necessary. Learning outcomes should be selected from the listbox on the form to accompany the assessment and indicates what the student is expected to achieve within the unit. Assessment criteria should relate to these learning outcomes. The assessment criteria are added to a treeview on the form. The designer can select whether to use sub-criteria and in this case sub-criteria will be required. Each criteria and sub-criteria has a title and a description. A weighting for each criteria within the assessment must be entered so that marks can be calculated correctly. Optionally learning outcomes may be selected for each individual criteria, although the designer may prefer to leave the learning outcomes for the overall assessment rather than attempt to be too fine grained in their interpretation within the individual criteria.

A listbox with the grades within the grading scheme is shown on the form. For each grade the description of requirements that the student has to meet in order to achieve that grade should be entered. Optional feedback for improvement may be entered for each grade too so that students can receive generic hints on how they could improve their work to achieve a better grade.

Assessments can be saved and loaded from file. Figure 3 shows how the report criteria were represented in the application along with supporting comments for each grade.

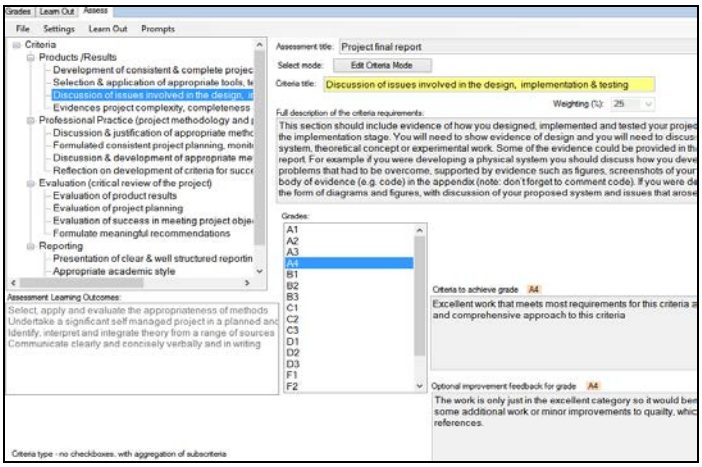

Figure 3: Report assessment criteria in the application.

4.0 Use of the Application in Marking

Once the assessment was configured within the application it was used to mark a small sample of project dissertations.
In marking mode a student name must be entered. For each criteria the assessor must make a selection from the grade listbox that corresponds to the appropriate description relevant to that element of their assessment. They may also add specific manually entered comments for each criteria. They can also add comments for the whole assessment of that student's work.

Once the work has been marked the student's marks can be saved and an overall mark calculated. A report on the student work can be produced that can be reviewed in another window. The report contains the description of the assessment and learning outcomes. For each criteria and sub-criteria the report can include a description of each, the grade awarded along with generic feedback for the grade, description of requirements to achieve the grade, and feedback for improvement, along with any manually entered comments. At the end of the report an overall grade and mark can be included along with overall comments for the whole work. This can be printed to give to the student or saved for later. Figure 4 shows a sample of the feedback report that could be given to students with a representation of some of the feedback that can be generated for the project report.

The report may take the form of a multi-page document, however too much detail may confuse students and detract from the usefulness of the information contained in it so that students may not read it. Therefore the assessor can choose which elements of feedback can be included in the report from a pop-up window. This will vary the length and detail contained in the report.

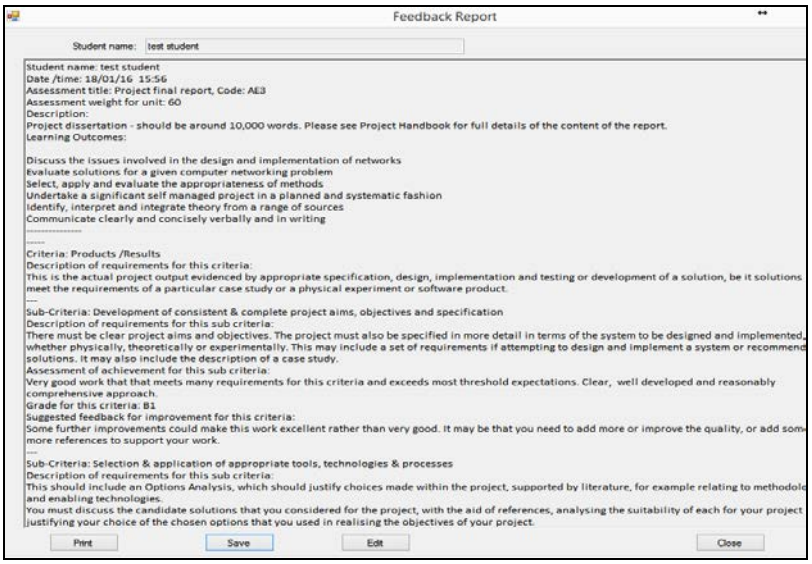

Figure 4: Sample of Feedback report for students.

5.0 Conclusions and Recommendations

The application was tested by marking 18 final year project reports, half for first marking and the other for second marking. The reports were marked online using a single monitor with both the report and prototype application open at the same time. Use of the application in this way seemed to increase the speed of marking compared to previous methods and yet it was able to generate more detailed feedback. Feedback was made available for students, although at this stage they were unable to make use of it for improvement. The response from students was not measured at this stage. Nevertheless the tests involving this case study demonstrated that the aims of the application had been achieved.

The next stage will be to use the application at an earlier stage of the project, on other units and involving more of the course team. As one of the principle aims of the application is to provide timely feedback to students their response to this should be measured. One method will be to design an appropriate questionnaire. Students can also be encouraged to use the application for self-evaluation.

6.0 References

1. Wiggins GP. *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco, CA: Jossey-Bass Inc.,U.S.; 1998.
2. Palmer N, Batola J. Strategies and experience in dealing with widening participation and diversity in teaching computer networking. *HEA Conference*. Nottingham: HEA; 2006.
3. HEA. *A marked improvement transforming assessment in higher education*. HEA, 2012. Available from: https://www.heacademy.ac.uk/sites/default/files/a_marked_improvement.pdf [Accessed: 22nd January 2016]
4. Stevens DD, Levi AJ. *Introduction to Rubrics: An assessment tool to save grading time, convey effective feedback and promote student learning*. United States: Stylus Publishing; 2012
5. Palmer N. Work in progress - Automation of a computer networking laboratory. *Global Engineering Education Conference (EDUCON), 2015 IEEE*. Tallinn: IEEE; 2015.
6. Nagel C, Evjen B, Glynn J, Watson K, Skinner M. *Professional C# 2008*. Indianapolis, IN: Wiley, John & Sons; 2008.
7. Sadler DR. Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*. Routledge UK; 2005; Vol 30(2): 175–194.

8. Stevens DD, Levi AJ. *Introduction to Rubrics: An assessment tool to save grading time, convey effective feedback and promote student learning*. United States: Stylus Publishing; 2012.