

Analysing and Improving Mathematical Formulation of Webavalia: A Self and Peer Assessment Tool

Rosalina Babo
Information Systems Department
ISCAP/ P.Porto
Portugal
babo@iscap.ipp.pt

Ricardo Fitas
Mechanical Engineering
FEUP, University of Porto
Portugal
up201704308@fe.up.pt

Jarkko Suhonen
School of Computing,
University of Eastern Finland
Finland
jarkko.suhonen@uef.fi

Markku Tukiainen
School of Computing
University of Eastern Finland
Finland
markku.tukiainen@uef.fi

Abstract—Recently, Problem-based learning has been increasingly used, and with those practices, self and peer evaluation tasks have gained more importance. This method of learning assists learners in the acquisition of skills and competencies. In order to support self and peer evaluation tasks, there are some different software offers in the market. The ease use of them is demanded since these tasks are an important part of evaluation. With the literature emphasis on the importance of assessment and critical spirit, it arises the possibility to implement formulas to increase the efficiency on the methodology applied and satisfaction of students and teachers. This paper introduces formulation problems about marks weighting in self and peer evaluation. It presents the advantages and disadvantages of different solutions for the formulation and suggestions of adapting them among tools, feedback surveys, and teacher’s methods of evaluation. It is also expected that the formulation, which was implemented in the software tool WEBAVALIA, has a great impact on feedback, improving user experience, and both efficiency and flexibility.

Keywords—mathematical models, marks, self-evaluation, peer evaluation, e-assessment, algorithm optimization

I. INTRODUCTION

Problem-Based Learning (PBL) is being more commonly implemented in higher education, since group working can improve important skills such as creative thinking and problem-solving, as well as the learning outcomes [1]. With the increased use of PBL, self and peer evaluation have also gained great importance in individual and collective schoolwork. During the evaluation, providing feedback is vital to improve the long-term practice [2] and the project work that is commonly part of PBL. As most projects require the formation of groups, peer evaluation gains more significance, because group members may not have the same performance during the project’s development [3]. Therefore, in general terms, students’ feedback emerges as a crucial part of the workgroup, and the reason why so many authors, such as Frehner (2019) [4] and Dochy et al. (1999) [5], defend the weight of self and peer evaluation on the project’s final mark. Dochy et al. [5] said that if students are involved in the project assessment, it can be “perceived as being valid, reliable, fair and as contributing to a growth in competence.”

Despite the advantages and the importance of students’ assessment, literature also defines disadvantages. For example, the amount of time necessary to perform the assessment; and the possibility of the students not differentiating the remaining members or jeopardising the

others. These aspects are important throughout the analysis on the inclusion of evaluation weights on final marks.

The improvement of time solutions is being published along with the evaluation of technology. Nowadays, there has been the development of software to assist with the assessment, in education, since it is necessary to implement easy ways for students to work in groups. Today, it is possible for students to do self and peer evaluation through evaluation tools. However, these tools need to be more efficient, with intuitive and easy tasks.

According to our experience, most of these tools use surveys to gather general feedback on the work developed. This approach solves most issues and time costs, because evaluation passes through data collection and analysis which can be very time consuming. However, with the importance of students’ peer evaluation and, at the same time, its disadvantages, it is crucial to do a serious analysis over the potential mistakes that students can do unconsciously when evaluating the others as, for example, evaluating all the group members with the same mark. When performing peer evaluation, there is always the risk of the students’ marks being jeopardised. At the same time, it is important that the students’ performance is assessed “properly and fairly” [6, p. 93], which may entail stronger penalties in some students. Therefore, it is essential to answer: how much does the student self-evaluation counts on their grade? What is the appropriate penalty over their mark?

At the moment, there are some implementations tested and reported in scientific papers, such as WebPA [7], Willey (2008) [8], PeerWise [9], and Cook (2017) [10]. All these examples incorporate algorithms that are implemented in the students’ evaluation such as a weight on their own marks, and all have different accepted methodologies. The present paper delivers formulation of another methodology used in the implementation of a different software – WebAVALIA, which has as its main goals being easy and free to use. WebAVALIA is a tool that aims to be more flexible with situations that frequently occur in workgroups and that cannot be predicted as, for example, the dropouts. Considering that the different tools should be studied in terms of efficiency and algorithmic optimisation, this study leans over four different formulas that were and are implemented in order to weight marks in the evaluation. The main parameters about the different issues on formulation will be referred in this paper. As authors know the possibility of a continuous improvement of the tool that is being implemented, this work counts with detailed discussion and future work, not only for this specific study, but for a global improvement of issues that still exist.

The present study introduces the background of concepts which reflects the main topics such as assessment, PBL method of evaluation, workgroups, its importance in developing and improving skills and competencies, as well as a small existing tools comparison. Section III has a brief presentation of WEBAVALIA. Then the design and development methodology of the tool is explained. Section V introduces a reflection on mathematical formulation considerations. Consequently, section VI explains the formulas implementation, a critical reflection on their impact, and some representative examples. Finally, the conclusions and future work are presented.

II. BACKGROUND OF CONCEPTS

Assessment is an important feature to assure the effectiveness of the learning practices. It can influence the learning outcomes, as well as the learning and teaching processes. The development of the technology led to the innovation of the assessment practices; therefore, e-assessment systems emerged to support the assessment process, by increasing its efficiency and reduce its costs [11]–[16].

The increased use of learning methods for the acquisition of skills and competencies resulted in the rise of Problem Based Learning (PBL). According to Hmelo-Silver [17], PBL is a method of learning that assists the learners to become active, by providing them with a problem that needs solving. This approach can be implemented with collaborative groups, which, in addition to the skills improved by PBL, also assists the students to achieve other important skills, such as autonomy, reasoning, critical thinking, communication, and problem-solving skills, as well as self-learning, peer-learning, accountability and team learning [18], [19].

However, workgroups can have some issues during the assessment of individuals. Therefore, the evaluator can implement self and peer evaluation. Self-evaluation involves the students in their own learning by actively participating in the assessment of their knowledge, performance, and learning [20], [21].

Peer evaluation is an activity in which individuals have the opportunity to consider their peers work, by making judgements on the quality or value of their performance. There are many variants of this process, but generally involves providing feedback to assist in the improvement of their peers' performance [22]–[24].

Besides improvements on the quality of their work, peer evaluation can also provide the development of other skills, such as collaboration and communication skills, improvement of learning quality and understanding, as well as problem-solving skills. These skills also allow the individuals to self-assess and judge their own performance in a more accurate way. For this reason, peer evaluation has proven to be effective in various contexts, from elementary to higher education [22]–[24].

Currently there is the use of technology to support assessment tasks, therefore the amount of self and peer assessment software tools existing in the market has been increasing. Each of these tools has their own features that can suit each evaluator's needs. Some of the existing tools found online are InteDashBoard, iPeer, Peergrade, PeerMark,

TeamMates, WebPA, Workshop Module, among others [25]–[31].

An analysis on these tools, show that most of these tools only perform peer assessment, being Peergrade, WebPA, and Workshop Module the exceptions capable of performing self and peer assessment. Some characteristics to take into account when analysing assessment tools are the type of questions presented by the tool, its capability to generate workgroups and/or assign members to groups, the possibility to have different assessment weights, and to restrict the evaluation given by the students, the tool pricing, among others.

When analysing these tools, and considering the previous characteristics, it is possible to perceive that each tool has their own question type such as rubrics, QA surveys, comments, and others. Every tool, with the exception of Peermark, can generate/assign workgroups. Peergrade and Workshop Module have the capability to perform weighted evaluations. While none are capable of restricting the evaluations given by the evaluated. When regards to availability, any of these tools are freeware.

III. WEBAVALIA

WebAVALIA (Fig. 1) is a self and peer evaluation tool to support the evaluator with the task of differentiating members in a team. The tool aims to provide an easy, quick, anonymous, and fair evaluation. There are two main users of this tool, the group member and the evaluator. Depending on their role, the tool enables certain features.

The group member is the one being assessed and can only use the voting board (Fig. 2) to evaluate him/herself and the remaining members of the group. This board registers the scores given by each member of the group in each assessment moments. The assessment moments are the times when the voting occurs.

The evaluator is responsible for the assessment, therefore has access to most features. The evaluator can create editions where the assessment occurs. When creating an edition, the evaluator has to set the parameters, assign members to groups, give the project marks, and calculate the results from the scores provided by the members.

The configurable parameters (Fig. 3) are the weights of the self-evaluation and the peer evaluation, which can be from 1 to 5. The number of assessment moments (1 to 3) can also be set, as well as the respective weights of each moments, that can be established from 1 to 5.

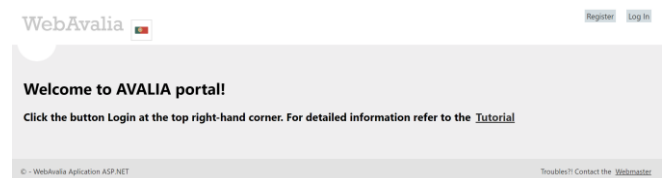


Fig. 1 WebAVALIA

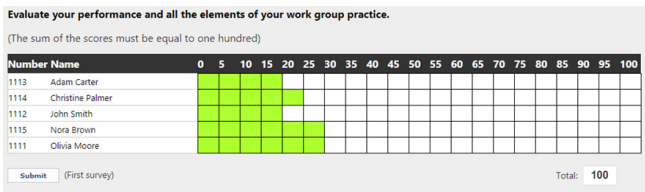


Fig. 2 WebAVALIA voting board

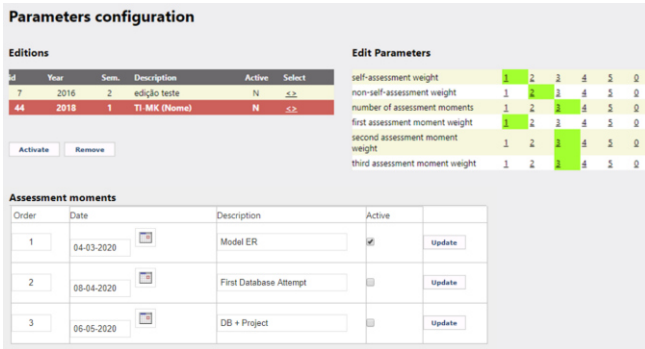


Fig. 3 WebAVALIA configurable parameters

IV. METHODOLOGY

WebAVALIA was developed using Design Science Research (DSR) methodology. DSR methodology is an iterative three cycle (relevance, rigour, and design) research for the creation or improvement of solutions to relevant problems. Therefore, WebAVALIA was developed to provide support with the task of distinguishing the students' performance in a workgroup assessment. It was built in ISCAP and has, as its knowledge foundations, the concepts of assessment, self and peer evaluation, workgroup assignments, PBL, and other e-assessment software tools offered in the market [32]–[35].

The design cycle encompasses all cycles in order to develop the framework. After its development, WebAVALIA had to be evaluated in order to identify weaknesses and possible improvements. From this evaluation, alternatives were made and then implemented. The evaluation and redesign iterative feature of this methodology ensures that WebAVALIA will be in constant improvement and development [32]–[35].

WebAVALIA development was prompted by the need to distinguish students in a workgroup. In the beginning, there was not much thought about a final result, the tool was just being evaluated to achieve a better solution. Nonetheless, it was being evaluated and feedback was gathered, and one of the weaknesses identified in the evaluation phases concerned the individuals marks, which were provided by mathematical formulations. Therefore, by improving the results given by the tool, different formulas were developed to better distinguish and assess the students.

This article provides the evolution of these formulas, as well as a comparison between them, by discussing their flaws and strenghts.

V. CONSIDERATIONS ON FORMULATION

In this section, considerations to take into account when developing a self and peer evaluation tool will be

enumerated. Another important aspect to consider is the several types of evaluation in a group or between groups, such as self and peer evaluation, and project evaluation.

Mathematical formulation is a common form of expressing, in terms of quantity, models for description of phenomenon and is well understood worldwide. Similarly, mathematical weights and formulas are used to classify a student by his own performance. It exists to create fairness between the work developed by the student and the expected learning outcomes. With this, the teacher perception and the relative perceptions of the student in relation to the others' performance creates an important structure to formulate.

Specifically, in self and peer evaluation, it is possible to detail and explain the different parameters into the structure of the students' project marks:

1. Distinguish between the most and the least engaged student

Self and peer evaluation can be advantageous, because it can be an opportunity for students to think about the global performance of their group and to improve critical thinking and evaluation skills [7]. These skills can be used to distinguish students in the same group at an individual level [36]. This distinction promotes students' engagement on the project, because they will have better marks when they are more involved in it. Therefore, this can be a motivation for students to work hard.

2. Distinguish between the work of the different groups

Besides the distinction between individuals in the same group, it is also necessary to distinguish the different groups. Usually, these differences are parameterised with the project mark classified by the teacher. This means that when two groups present their work, depending on the performance of each individual of the group, the groups between them can have different marks. Peer evaluation can provide a fairly distinction between the several groups to determine the best group.

3. Differentiate between the best and the worst students in terms of expected contribution

Other factors on students' perspectives have to be consider, for example the different evaluation goals of the students. Some students may want to achieve a higher mark, while others have a different perspective and only want to achieve a passing grade. There are also other cases where the availability of the students may not be the same, depending on their personal schedule. If these students are assigned randomly to the same group, a conflict of interests may exist. When a student that has a busy schedule is assigned to a group with an ambitious student, the later will have to work harder to achieve his/her goal, since the first has limited time to dedicate to the work. This issue can be easily solved by dividing initially the groups according to the students' preferences. This method can be empirically used, although there may exist students that have no idea about their colleagues' interests or schedules, which can create difficulties when assigning individuals to groups.

When a group is formed, it normally remains the same for the rest of the sequential evaluations. Nonetheless, there may be members that change groups due to conflicts, as previously explained. When students change projects, it can

be difficult for them to adapt. Also, the teacher will have more difficulties in the assessment process, and therefore these changes should only be done before the first assessment.

Another perspective to considering when evaluating a group project is the students' knowledge background and familiarity with the concepts. A student who never had contact with the key concepts of the project will have to work hard and endeavour to achieve the expected learning outcomes. While a student who has previous knowledge on the subject will not have to give much personal effort to comprehend the topics. Nonetheless, the first student, even though having achieved all the learning outcomes, may not have an equal or better mark than the second student, despite his/her efforts. This may happen because the student with background knowledge was more at ease with the concepts, and only ended up developing his/her understanding on said concepts. Therefore, the effort and commitment on the project development should be considered when assessing the individuals in a group.

Considering what was just discussed, this can also happen when evaluating individuals in a group. For example, a student who does not have familiarity with the project topics is assigned to a group where the remaining members have full understanding about the concepts. This student may not have to endeavour to achieve a higher project grade, because his/her colleagues understand the goals of the project and can easily accomplish it. Consequently, even though the student did not attain all the learning outcomes, the group project mark was high.

The opposite can also occur. In that same class, there is a hard-working student who is assigned to a group with members who have few competencies. The members of this group will have to work harder to achieve the same learning outcomes of the previous group, but nevertheless they may not obtain the same mark as the student of the other group, despite their endeavour.

Hence the importance of distinguish the students learning outcomes according to their expected contribution to the project.

4. Parameterise students' marks considering teachers' perceptions

When elaborating an assessment that consists in the development of a project, the evaluation must have an effective method to distinguish students. The teacher must also, previously, develop an evaluation grid in order to be possible different teachers achieve the same marks in the same groups. Therefore, a framework capable of calculating the students' self and peer evaluation of the project's learning outcomes is necessary. These features make the tool less rigid, and the assessment more flexible to the evaluator.

5. Consider the dropout students.

The workgroup assessment process can be even more complex when a student drops out, because this student may be present in class, but not working in a group. Unless, the student expresses the dropout, the teacher may not know, and the remaining members of the group are overloaded and can be jeopardise.

All the examples taken of the literature try to attribute a mark based on the first two variables listed: the distinction of the students, at an individual level and at a group level. However, some students have their own preferences among the methods each teacher implements, and the teachers have their own necessity to implement a different method. Also, the students' grades can be impacted if a group member has to leave the project for any external reason. Therefore, it is important to have formulas that can circumvent these variables.

These issues can be solved with a software tool called WEBAVALIA. WEBAVALIA is a free and simple software tool that aims to facilitate the students' self and peer assessment process and the teachers' task of assigning marks. The necessity to solve most of the disadvantages on the methodology used in other papers and to implement, test, and improve the tool in terms of accuracy and user experience (UX) allowed the creation of WEBAVALIA. The methodology used on this tool is inspired on the above-mentioned variables, using formulas that respect them. These formulas were also improved based on students and teachers' feedback.

Some characteristics on WEBAVALIA may differ in relation to other tools, including the main objective and the assessment and weighting methods. For example, Peergrade [37] is a powerful tool capable to do assignments reviews. It can also do automatic statistics for a more precision evaluation, however, due to the number of features implemented, most of them are unlikely to be used. Other tool is WebPA [6], that can be flexible by allowing teachers to make their questions. It is possible to make the assessment with no restrictions; therefore, the students are able to evaluate all their peers with the same mark. Consequently, if it occurs, no conclusions can be found to differentiate the group members, which is not an optimal solution. These examples show the motivation in improving mathematical models of WEBAVALIA, due to the simplicity and, at the same time, better accuracy on solving the complexity of the referred issues.

Therefore, the following section will present the formulas which were used in order to sustain the topics presented and were improved continuously.

VI. ANALYSIS AND COMPARISON OF FORMULAS IMPLEMENTED IN WEBAVALIA

Notwithstanding the discussion in the previous section on the individual endeavour, methods of comparison between different projects results, and the application of different assessment criteria in the same project, and upon overcoming these difficulties, it is also important to understand the need to distinguish the performance of each individual in a group. The evaluator, when using WebAVALIA to obtain final marks, has to already have a comparison between the different groups and a project grade. Therefore, WebAVALIA can only be used to distinguish the students in the same group.

Consequently, the quantification of the group members performance perception is important to reflect the students' and teachers' points of view. Most of the evaluation is scored with numbers, due to the quantification aspect of assessment, normally given by the marks and scales, it was important to develop mathematical formulas in WEBAVALIA. This tool

has some advantages, such as being free, flexible and simple, which facilitates the voting process, making the self and peer evaluation easier.

WebAVALIA configurable parameters are the weights of the self-evaluation (w_A) and the peer evaluation (w_B), which can be from 1 to 5. The number of assessment moments (1 to 3) can also be set, as well as the respective weights of each moment, that can be established from 1 to 5. These weights are parameterised by the teacher and in order to simplify the formula presentation, only an assessment moment was considered. The application of these weights is done in the final score calculations of self and peer evaluation, expressed by M_k .

All the students of the group (S_1, S_2, \dots, S_n) assign the respective mark for self and peer evaluation. Let us assume A_k (between 0 and 1) as the mark of the self-evaluation of the student k and $B_i, \forall i \neq k \wedge i \leq n$ as the peer evaluation marks. Some weights, w_A and w_b , are adjusted by the teacher in order to allow them to penalise more the self-evaluation in relation to peer evaluation. Where one of the students will take an average mark, as follows (1):

$$M_k = \frac{w_A * A_k + w_B * \sum_{i=1, i \neq k}^n B_i}{w_A + w_B * (n-1)} \quad (1)$$

First, it is important to establish that there are some conditions that do not allow the students to assign a full mark on both self and peer evaluation. In most tools, this can occur, which means that it may not be possible for the teacher to have a real perspective on the actual performance of the students. WEBAVALIA considers this condition, which allows a better distinction of the students. The inexistence of this condition, in the limit, would cause a non-distinction of the group members. The condition implemented in WebAVALIA for A and B marks are that:

$$A_k + \sum_{i=1, i \neq k}^n B_i = 1 \quad (2)$$

An example of the implemented formulas is presented. A group is formed by 4 students, John, Peter, Anne, and Susan. They undertook 3 assessment moments with the same importance, meaning the same weight, along the project and the results of their self and peer evaluation are presented, in percentage, in Table 1. The values used for w_A and w_B were 1 and 3, respectively. It is also important to note that these values represent the self-evaluation, which means that the self-evaluation is considered three times less relevant than the peer evaluation, that is a portion of 25%, in this case.

TABLE I. STUDENTS' SELF AND PEER ASSESSMENT MARKS

Evaluators	Evaluated				M_i
	John	Peter	Anne	Susan	
John	28	30	28	32	30
Peter	13	17	15	18	16
Anne	22	23	22	23	23
Susan	37	30	35	27	33

The table shows that Susan has the highest mark of 0.33 out of 1. John has a contribution of 0.3. Anne and Peter have the lowest scores, being Peter, the student with the lowest score in the group. Curiously, the sum of all evaluations should be 1, but the sum of the values of the last column are equal to 102, that means that the calculations were approximated to be integers, so there is 2% of approximation errors. As the following calculations are always compared with the maximum value or with the average value, this error will not contribute for the error of the calculation of the final mark.

Comparing the condition of (2) with the existence of no condition, it is possible to realise that instead of each student evaluating themselves and their colleagues with 1, they evaluated with $1/n$, thus the result is numerically the same and independent of the values of w_A and w_B . However, it would be more realist if all members performed the work equally between them, since all did a certain percentage of the whole work.

With all marks M_1, \dots, M_n calculated, let us assume M as equal to the maximum value of this set, and continue with M_k as the mark of self and peer evaluation of the students. Also, it should be important the insertion of a variable with reference of teacher's evaluation, referent to parameter 2 of the section V. This variable C is set as being the teacher's classification of the project.

The first formula implemented was Form. A. This formula assumes that exists a linearity correlation of the final mark in relation to mark M_k . This also guarantees that the students with $M_k=0$ and those which $M_k=M$ have a final mark of C_k equal to 0 and C , respectively. That means the formula is given by (3):

$$C_k = \frac{M_k * C}{M} \quad (3)$$

This formula was tested, and it indicated that the final marks had a broad range, this means that this formula caused a huge difference between the students' marks. This is because half of the range of possible marks can be a cause of fail (0 to 10), while an acceptable mark range varies from 10 to 20. If this method is applied in a course where the project is equal to the final grade, and if it has various assessment moments, it will benefit the students who work better between each moment, but highly jeopardise the students who work less, in order for the evaluation being almost approached by $1/n$ for each one of the individuals. There may also exist some cases, where the student jeopardises another student by assigning a low score, causing the failing of said student in the course. This formula considers that if the first assessment moment is done, the groups should be permanently the same. Another point of view can be assuming $C=20$ and having an interval with a range of 0.01, which can mean that the worst student has 0.245 and the best has 0.255. Despite the grades being almost equal, the value of C_k when M_k is the minimum of range, equals 19, which is not fair.

Some simple applications of this formula are the cases where the teacher has a weight for the project mark and another weight for the student's evaluation score. A downside of this formula is that the mark is not 0 for the students who did not work.

Another formula (Form. B) was tried, but without success, due to some incorrections on the formulation. As a formula C was being developed, formula B was not considered for implementation. However, the authors present in this paper an alternative for the formula, with the following explanation.

Based on formula A, this formula (3) consists in increasing the final mark for those whose mark is less than the average of all self and peer evaluation marks. Let it assume \bar{m} as the arithmetic average of the set of students' marks. Therefore, instead of having a linear proportion, it has two linear proportions and a constant (4):

$$C_k = \begin{cases} \frac{M_k * C}{M} \leftarrow M_k > \bar{m} \\ \frac{\bar{m} * C}{M} \leftarrow \frac{-2}{M} \leq M_k \leq \bar{m} \\ \frac{M_k * C}{\bar{m}} \leftarrow \text{otherwise} \end{cases} \quad (4)$$

The advantage of using this formula in relation to Form. A is because with M_k value lower than the average, the worst students do not get too much jeopardised. The 2nd equation of the system results of the transition between the worst and the best value of each linear functions. The disadvantage remains in the non-flexibility, since there is not a manner for the teacher to change the assessment range between the lower and higher marks in a group. Also, the students are still compromised due to the method. If the range in Form. A is increased to 0.02, the marks' difference will have no change, even when the mark is increased by 0.02. In formula B, in a difference of 0.01, the marks will be the same, however by raising 0.02, the marks will be different. Nonetheless, this insignificant difference between students' votes will result in a considerable distinction between student's final marks.

There was the development of a formula C in order to solve the problem of flexibility, and to shorter the range of marks, minimising the existence of low marks. Considering the introduction of a parameter $T \in \{1, 1.25, 1.5, 1.75, 2\}$, the equation (5) will vary also in function of T, allowing the teacher to boarder of shorter the marks' range of the students.

$$C_k = C - \left[\frac{M - M_k * C}{M} * \frac{C}{T} \right] \quad (5)$$

In fact, respecting parameter 4, it is possible to change the value of T if the main objective is to increase the range of values in function of the dispersion of students' self and peer evaluation marks. This is why, Form C is better than Form B, because variable T allows to increase or decrease the range to achieve fairer results. However, when using this formula, in the case when the student has zero with T=1, that student will have 0 as a final mark. However, when T=2, that same student will have a mark, when that should not happen, while the remaining members of the group will be jeopardise if T=1, which is unfair, since these students worked.

The main objective of the third formula (Form. C) implementation is that the teacher can give a value to a parameter in order to modulate students' marks, allowing a higher or lower range marks.

In order to better consider the value of the highest scores, it is important to assure it exists a score that the groups consider to be the lowest mark to give. Therefore, if the rest of the students should be marked with as much as the best student, this behaviour can be optimised by a formula D (9),

$$C_k = C * \sqrt{\frac{M_k}{M_s}} \quad (6)$$

where $M_s=M$.

Considering the example above (Table 1), when the project mark provided by the teacher is 17/20. Table 2 presents the calculations of the final marks, according to the formulas A, B, C, and D.

TABLE II. STUDENTS' FINAL MARKS, ACCORDING TO FORMULAS A, B, C, AND D

Students	Form. A		Form. B		Form. C		Form. D
	C_i	Int. C_i	C_i	Int. C_i	T=1	T=2	C_i
<i>John</i>	15.4	15	15.4	15	16	17	16
<i>Peter</i>	8.2	8	10.9	11	9	13	12
<i>Anne</i>	11.8	11	13.0	13	12	15	14
<i>Susan</i>	17.0	17	17.0	17	17	17	17

As shown in table 2, it is evident that by using formulas A and C, and considering that the project's passing mark is 10/20, Peter fails the project. The expected value for each individual is $\frac{1}{4}=25\%$ and Peter had a voting average of $C_2=16$, meaning he achieved 16 points, given by the self and peer evaluation weight, which represents 64% of the expectation. However, it can be seen in another perspective, by considering the hypotheses of Peter being jeopardise by his colleagues, therefore the expectation would be a little higher than the previous value. In this case, Peter would achieve a higher mark, which can make the difference between achieving the minimum mark to succeed in the project. Even if the project was evaluated with a 17/20, and in consequence Peter was probably very compromised by the peer evaluation. Consequently, there are many points of view regarding this problem, then it is important that the teacher has a variable T to adjust Peter's mark.

In addition, it is possible to observe a great dispersion of the results among the students with the lowest marks of the group. Figure 4 presents a graph where it is possible to observe the dispersion. Note that the values were transformed from \mathbb{Q} to \mathbb{R} set, since the representation is clearer. It should be possible to use parameter T for formula 4 in order to allow the teacher to model the best grades for the students. Also, this formula can be used more efficiently, since it is also coherent with parameter 5.

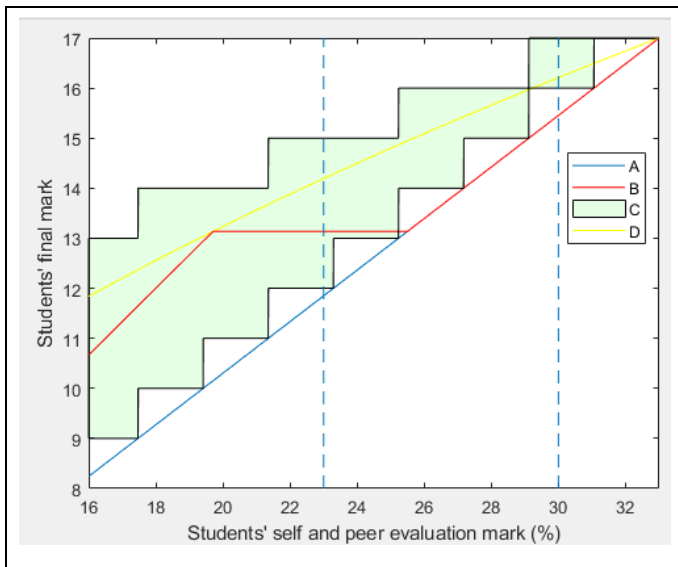


Fig. 4 Graphic representing formulas A, B, C, and D.

VII. CONCLUSIONS AND FUTURE WORK

The present paper presented the self and peer assessment tool – WebAVALIA, as well as regarding the design cycle of DSR by explaining the mathematical formulations considering the different parameters and the gathered feedback. It intended to detail the issues related to the weighting of students' marks on the evaluation of their own and their peers' performance in a workgroup, compare different implementations on literature, and introducing new methods, in order to extend the studies on this theme, which is transversal throughout all areas, including academic, business organisations, or sports teams [38]. This exposal of critical thinking about implementing formulation and how it can be damaging for the students if they are not well designed is important.

Four different formulas (A, B, C, D) were developed. Each formula has some downsides that were identified and led to the motivation to design others. Formula D is free of user dependence, despite it being the most acceptable manner of weighting, among the proposed, it is important to make the formulas more complex, in order to attain an optimal method of satisfying all the students' expectation as much as possible. While it is important to maintain the simplicity necessary to develop a simple and powerful tool, it is also important to allow the students to express their feedback on the remaining group members.

For future work, the goal is to develop new implementations to reach the most beneficial outcomes for every student and teacher, in order to make the tool more flexible and efficient for every type of workgroup. Since there is not a standard grading scale to mark tests or assignments in educational institutions, one of the future implementations might be the capability to adjust different marks scales (e.g. 1 to 10; 1 to 20; 1 to 5; etc) in order to suit other evaluators' needs. These implementations will help to achieve a better perception of fairness to most users, as well as provide fair results for all borderline cases.

It was implemented in WebAVALIA, in Form. D, a fixed value of $M_s = M$. A future development of Form. D can be to make M_s dependent from all the scores of the group, which will cause dependence between scores, achieving fairness

since the final result will not depend on the best score but all the relative scores. Similar to Form. C, it can be implemented a T variable, making the formula more flexible.

It would be interesting to study WebAVALIA's learning process and usability by the teachers in order to understand the difficulty in learning and using this tool. Also, an analysis on the students' opinions and experience using WebAVALIA will be discussed in a future study. This analysis is performed taking into consideration the students' feedback. At the end of the academic year, all the students who have used this tool for self and peer assessment are invited to answer a survey where their opinion is registered to future analysis.

Nowadays, the implementation of new methods of Artificial Intelligence (AI) are being used in almost all areas of study. Regarding the improvement of WEBAVALIA, it could be interesting to implement an iterative formulation that is updated with the student's feedback on the assessment tool. There are already some studies about this topic, but there is not a specific tool that implements the methods on higher education assessment.

ACKNOWLEDGMENT

We would like to express our gratitude to Vítor Silva (MEng, software developer) for his important assistance with the development of AVALIA/WeAVALIA tool. We would also like to thank the graduate student Joana Rocha for her collaboration in writing assistance, language editing, proofreading support, and software testing.

REFERENCES

- [1] A. J. Khoiriyah and H. Husamah, "Problem-based learning: Creative thinking skills, problem-solving skills, and learning outcome of seventh grade students," *Jurnal Pendidikan Biologi Indonesia*, vol. 4, no. 2, pp. 151–160, Jul. 2018.
- [2] M. Frank and A. Barzilai, "Integrating alternative assessment in a project-based learning course for pre-service science and technology teachers," *Assessment & Evaluation in Higher Education*, no. 1, pp. 41–61, 2004.
- [3] D. Hall and S. Buzwell, "The problem of free-riding in group projects: Looking beyond social loafing as reason for non-contribution," *Active Learning in Higher Education*, vol. 14, no. 1, pp. 37–49, 2013.
- [4] M. Frehner, "Self- and peer-evaluation of individual project work: An innovative course assessment method to increase student motivation," 2019, pp. 5–41.
- [5] D. Sluijsmans, M. Segers, and F. Dochy, "The use of self-, peer and co-assessment in higher education: A review," *Studies in Higher Education*, vol. 24, no. 3, pp. 331–350, 1999.
- [6] N. Hosseinzadeh and F. Mnif, "Assessment of individual students in group work in engineering programs," in *IEEE Global Engineering Education Conference, EDUCON*, 2016, vol. 10-13-April-2016, pp. 93–99.
- [7] S. Loddington, K. Pond, N. Wilkinson, and P. Willmot, "A case study of the development of WebPA: An online peer-moderated marking tool," *British Journal of Educational Technology*, vol. 40, no. 2, pp. 329–341, Mar. 2009.
- [8] K. Willey and A. Gardner, "Improvements in the self and peer assessment tool SPARK: Do they improve learning outcomes?," *Proceedings of ATN Assessment 08: Engaging Students with Assessment*, no. 1, 2008.
- [9] "PeerWise." [Online]. Available: <https://peerwise.cs.auckland.ac.nz/>. [Accessed: 17-Dec-2019].
- [10] A. R. Cook *et al.*, "Using peer review to distribute group work marks equitably between medical students," *BMC Medical Education*, vol. 17, no. 1, pp. 1–9, 2017.
- [11] M. Bahar and M. Asil, "Attitude towards e-assessment: influence

- of gender, computer usage and level of education,” *Open Learning: The Journal of Open, Distance and e-Learning*, vol. 33, no. 3, pp. 221–237, 2018.
- [12] N. A. Buzzetto-More and A. J. Alade, “Best practices in e-assessment,” *JITE-Research*, vol. 5, no. 1, 2006.
- [13] S. Kanwar, “Assessment - An important facet of learning,” *Literacy Information and Computer Education Journal*, vol. Special 1, no. 2, pp. 916–922, 2012.
- [14] A. Okada *et al.*, “Pedagogical approaches for e-assessment with authentication and authorship verification in Higher Education,” *British Journal of Educational Technology*, vol. 50, no. 6, pp. 3264–3282, 2019.
- [15] D. Pereira, M. A. Flores, and L. Niklasson, “Assessment revisited: a review of research in Assessment and Evaluation in Higher Education,” *Assessment & Evaluation in Higher Education*, vol. 41, no. 7, pp. 1008–1032, 2015.
- [16] D. Wiliam, “Assessment for learning: meeting the challenge of implementation,” *Assessment in Education: Principles, Policy & Practice*, vol. 25, no. 6, pp. 682–685, 2018.
- [17] C. E. Hmelo-Silver, “Problem-Based Learning: What and how do students learn?,” *Educational Psychology Review*, vol. 16, no. 3, pp. 235–266, 2004.
- [18] T. M. Daba, S. J. Ejersa, and S. Aliyi, “Student perception on group work and group assignments in classroom teaching: The case of Bule Hora university second year biology students, South Ethiopia: An action research,” *Educational Research and Reviews*, vol. 12, no. 17, pp. 860–866, 2017.
- [19] L. Wen, “The perceptions of students on cooperative learning at intermediate accounting II course,” *Business Education Innovation Journal*, vol. 9, no. 2, pp. 127–133, 2017.
- [20] S. Borg and A. Edmett, “Developing a self-assessment tool for English language teachers,” *Language Teaching Research*, vol. 23, no. 5, pp. 655–679, Sep. 2019.
- [21] F. Dochy, M. Segers, and D. Sluijsmans, “The use of self-, peer and co-assessment in higher education: A review,” *Studies in Higher Education*, vol. 24, no. 3, pp. 331–350, Jan. 1999.
- [22] K. J. Topping, “Peer assessment,” *Theory into Practice*, vol. 48, no. 1, pp. 20–27, Jan. 2009.
- [23] D. Reinholz, “The assessment cycle: a model for learning through peer assessment,” *Assessment & Evaluation in Higher Education*, vol. 41, no. 2, pp. 301–315, Feb. 2016.
- [24] L. Li, “The role of anonymity in peer assessment,” *Assessment & Evaluation in Higher Education*, vol. 42, no. 4, pp. 645–656, 2017.
- [25] “InteDashboard™ - Empowering TBL with Technology | Peer Evaluation.” [Online]. Available: <https://www.intedashboard.com/features/peer-evaluation/>. [Accessed: 14-Jan-2019].
- [26] “iPeer | Teaching with Technology.” [Online]. Available: <https://lthub.ubc.ca/guides/ipeer/>. [Accessed: 14-Jan-2019].
- [27] “Peergrade - engaging student peer review.” [Online]. Available: https://www.peergrade.io/?nab=1&utm_referrer=https%3A%2F%2Fwww.google.pt%2F. [Accessed: 14-Jan-2019].
- [28] “PeerMark™ - Guides.turnitin.com.” [Online]. Available: [https://guides.turnitin.com/01_Manuals_and_Guides/Instructor_Guides/Turnitin_Classic_\(Deprecated\)/23_PeerMark](https://guides.turnitin.com/01_Manuals_and_Guides/Instructor_Guides/Turnitin_Classic_(Deprecated)/23_PeerMark). [Accessed: 14-Jan-2019].
- [29] “TEAMMATES - Online Peer Feedback/Evaluation System for Student Team Projects.” [Online]. Available: <http://teammatesv4.appspot.com/>. [Accessed: 14-Jan-2019].
- [30] “WebPA.” [Online]. Available: <http://webpaproject.lboro.ac.uk/>. [Accessed: 14-Jan-2019].
- [31] “Workshop activity - MoodleDocs.” [Online]. Available: https://docs.moodle.org/35/en/Workshop_activity. [Accessed: 14-Jan-2019].
- [32] A. Hevner, “A three cycle view of design science research,” *Scandinavian Journal of Information Systems*, vol. 19, no. 2, 2007.
- [33] A. Hevner, S. March, J. Park, and S. Ram, “Design science in Information Systems research,” 2004.
- [34] A. R. Hevner and S. T. March, “IT systems perspectives - the information systems research cycle,” *Computer*, vol. 36, no. 11, pp. 111–113, 2003.
- [35] S. March and V. Storey, “Design science in the Information Systems discipline: An introduction to the special issue on Design Science Research,” *Management Information Systems Quarterly*, vol. 32, no. 4, 2008.
- [36] K. Pond, D. Coates, and O. Palermo, “Student experiences of peer review marking of team projects,” *International Journal of Management Education*, vol. 6, Jan. 2007.
- [37] L. IParadigms, “PeerMark™ Manual,” *Turnitin Instructor Handbook*, 2010. .
- [38] M. Stănescu and N. Mușat, “Quality analysis model of the e-learning training system for sports occupations,” *Procedia - Social and Behavioral Sciences*, vol. 180, no. November 2014, pp. 1351–1356, 2015.