

A Personalized Summative Model based on Learner's Effort

<https://doi.org/10.3991/ijet.v12i06.7165>

David Bañeres
Open University of Catalonia, Barcelona, Spain
dbaneres@uoc.edu

Abstract—Nowadays, instructors apply a large variety of learning methodologies to help learners to achieve the learning outcomes and to assess the knowledge acquired across the course. Formative and summative assessment models are mainly applied in multiple combinations independently of the learning environment (on-site, online or blended). When we move to an adaptive learning, the adaption tends to be in the learning process (learning path, activities, educational resources) mainly related to formative models but little adaption can be found related to summative models and very restrictive. In the latter case, grade formulas depending on performed assessment activities are typically defined to provide a personalized learning process. In this paper, we introduce the basis of an innovative personalized summative model based on learner's preferences and effort. Although this model conceptually may allow passing a course without evaluating all learning outcomes, it is not far from conventional summative models where a certain grade is required to pass the course and the learner may not have acquired all the knowledge taught in the course. The paper introduces the model and it also analyses an opinion survey on instructors and learners.

Keywords—adaptive assessment, summative model, personalization, learner's preferences, student's effort

1 Introduction

In the European High Education Area (EHEA), the instructional process is recommended to be based on a continuous assessment model: Students learn the different topics through all the course, and they are assessed with different learning activities and not only with a final exam. This methodology contributes to better acquire the learning outcomes since there is a continuous learning process during the course.

The previous statement is the recommended model and it can be achieved by following a formative model. In this type of models, learning outcomes, skills and competences are continuously evaluated to assess the knowledge acquired by the student and, therefore, continuous feedback can be provided to improve student's knowledge. For instance, rubric-based assessment [1] can be used to obtain a score of an activity based on learning outcomes.

However, most of the applied formative assessment models are combined with a summative model due to the complexity to evaluate a learner only by competences acquisition and to include some final validation (i.e. final exam). In this hybrid models, the summative model aims to grade students by adding several assessment activities during the course and a final validation, and the formative model helps to improve students' knowledge by providing continuous feedback based on their ongoing performance.

In this paper, we focus on a novel standalone summative model based on student's preference and by considering student's effort. Note that, this model is compatible with any formative model to boost the instructional process. The motivation comes from online learning where there are virtual learning environments (VLE) that allow customization of the learning process, such as the interface presentation or the learning path [2][3][4]. The grading system tends to be hardly specified by instructors and very constrained regarding how to calculate the grade of the learner. Moreover, student's effort is commonly not considered since each learner needs different effort to acquire the same outcomes or competences. In those models, learners tend to adapt their learning progress based on how they will be evaluated through the course. The new approach proposes to flip the adaptation. The student is not conditioned by the summative model. By the contrary, the assessment model is adapted based on her own learning preferences and by considering the effort invested on the course. That is, each learner may have a different grading model based on her preferences and exerted effort.

The paper is organized as follows. Section 2 analyses previous work on assessments models and effort estimation. Section 3 focuses on the learning perspective of the model and Section 4 introduces the mathematical foundations of the Personalized Summative Model. Section 5 shows the data structure designed to store the knowledge representation of the model. Finally, Section 6 performs a qualitative analysis of the model by analyzing instructors' and learners' opinions, and conclusions and future work are presented in Section 7.

2 Related Work

In this section, previous work and a motivation example are presented. First, a brief literature review is performed on hybrid assessment models. Next, different methodologies to evaluate student's effort are described. Finally, the motivation example is presented.

2.1 Assessment Models

In the literature, we can find many case studies and methodologies that apply summative [5] and formative [6] assessment models. If we focus on hybrid models that combine both models, interesting approaches are found. Authors in [7] reviewed instructional design models and compared them to instructional strategies in online courses and proposed a hybrid model to integrate constructivist and objectivist approaches with the instructional design. Another model was proposed in [8] using

multiple-choice assessment. Here, authors designed a new formative model using multiple-choice activities commonly used for summative models. Additionally, they described the advantages of multiple-choice test on formative assessment. In [9], a model for K-12 education level was proposed based on CBAL (Cognitively Based Assessment of, for, and as Learning) theory. This hybrid model aims to assess what students have learned and facilitates the instructional process. Other models were also proposed for final project evaluation [10], gamification [11] and evaluation of technical skills [12].

There is some discussion related to summative models whether they can be used standalone to provide a successful learning experience [13][14]. Authors in [15][16] analyzed particular courses where a summative model was applied and they concluded that they could be used to estimate student's knowledge and performance properly.

As far as we know, few personalized summative models have been proposed. Personalization is tended to be applied in the learning process. Authors in [17] described a customizable model for undergraduate students and a similar model was proposed in [18] for learning programming. Even a generic functional model was proposed in [19] to adapt the assessment.

2.2 Student's effort and measurement

The model proposed in this paper allows model customization based on student's effort. That is, the level of customization depends on the progression of the student and the effort she exerts during the course.

Here, the first handicap of the model appears: How is the effort evaluated? The effort is not a standard measure and each learner needs different effort to acquire the same learning outcome. The effort has been analyzed widely in the past decades and different research studies have been performed.

In the University of Benedict College (Columbia, USA), a policy combining effort and knowledge assessment is applied from 2004 [20][21]. The summative model takes into account both measures (knowledge and effort grade) in different proportion on first (40%, 60%) and following students (60%,40%). The author concluded higher effort contributes to get better grades and better knowledge.

In [22], a study in grading preference was done in physical education. The experiment concluded that teachers mainly preferred students who showed high effort compared to students with excellent physical abilities.

Previous models seek to maintain the engagement of students through all the course. Maintaining continuous engagement reduces dropout and the time spent in the course contributes to acquire knowledge. Authors in [23] evaluated how self-regulation could also affect knowledge acquisition. They found evidences revealing a positive relationship between engagement, effort and classroom performance on self-regulated students.

Effort can be measured using different strategies as we describe next:

- Time spent studying (or time on task) is an easy estimation. Here, different evidences have been used to collect the effort: student's self-reports [24], the

period of the week (i.e. daily or weekend) [25], class attendance [26], consistent effort across semester [27].

- Time off task [28]. Inversely, this estimation evaluates the time spent on the activity when the student is thinking about other tasks. If a student has a high time off task in one activity, she will feel that the task requires less effort.
- Mental effort [29][30] is defined as the amount of capacity or workload used by a learner to achieve a learning outcome. Note that, this effort is unrelated to time on task since mental effort evaluates the perception of the learner. This effort is also typically measured through learner's reports.

As we can observe, all these measures are subjective. There is no standard evaluation scale. However, it can be used to create a custom scale that can help to assess student's effort.

Other validation experiments analyzed previous measurements with the acquisition of learning outcomes [31][32]. Significant results were found distinguishing different student's profiles depending on the goal objective (i.e. obtain the best score or pass the course with the least effort). Another relevant conclusion is that all previous estimations are directly related to the outcome acquisition.

The next section introduces a motivation example: A simple score-based summative model with a very limited customization.

2.3 Motivation example: Score-based summative model

In this section, we motivate the new summative model. Authors in [33] described a score-based assessment system that combined formative and summative assessment and it could be used to define a broad set of models based on the type of assessment activities deployed on the course and the percentage of the score assigned to each activity. For example, let us assume the next summative model:

$$FM = \text{MAXIMUM} (35\% CA + 35\% FP + 30\% FE, 50\% FP + 50\% FE) \quad (1)$$

where the final mark (FM) of the course is computed based on two different models. The first one takes into account the continuous assessment (CA) activities performed during the course, a final project (FP) during the last month of the course and a final exam (FE). The second model only takes into account the final project and the final exam for students that have not passed the continuous assessment activities.

In this model, the learner could decide in some way which learning model wants to follow: a more formative one which is aligned with the different competences learned on each unit of the course or a more summative one based on a final project and a final exam that evaluate all the knowledge acquired during the course.

This type of adaption on the assessment model is very constrained. However, this adaption can be enriched by increasing the number of possible customizations. In this paper, we propose a novel summative model where the learner can select which activities she is interested in performing. Note that, the learner will be able to avoid even a final exam depending on her preferences.

However, the learner should have some limitations in this customizable model and not all students should be able to select the same preferences. We have to preserve the principal objective of the course from an educational point of view, that is, to acquire the learning outcomes of the course. For this reason, we propose that the learner should select the available preferences depending on previous *effort* or prior knowledge proved during the course.

Previous to describe the foundations of the proposed model, we discuss the learning perspective of the model. That is, the items that can be customized and the evidences that can be used to extract the effort exerted by the student.

3 Learning Perspective

This section puts emphasis on the learning perspective of the model. The customization items can be anything related to the calculation of the grade of students. We have identified the next items and some justification for each one is also provided.

- To add extra points to assessment activities: This is a common customization used on courses. When the learner performs some extra activities during the course, she can be rewarded with extra points on the mark. On the learning perspective, we can assume that the learner has reinforced her knowledge through the additional exercises. Note that, a range of different points (i.e. 0,25, 0,5, 1, ...) can be defined depending on the learner's effort.
- To extend the submission date: The learner can extend on demand the deadline when an activity must be submitted. On the learning perspective, this extension enables the possibility to improve the activity, practice the skills, acquire the competences, and get a better grade.
- To avoid an assessment activity: This is a more serious item. The student can decide to avoid an assessment activity and, therefore, the grade calculation formula should be readjusted based on the set of performed activities. Note that, the item can be added to any activity, even the final exam. The effort cost should be adjusted depending on the criticality of the activity in the assessment model. On the learning perspective, the learner should prove previous to the activity that she has acquired the knowledge addressed on the avoided activity.
- To avoid an exercise within an activity: This item is a particular example of the previous one. On complex activities, the learner could decide to avoid a particular question. The effort cost should be significantly less compared to the previous one and, in this case, only the grade of the activity should be adjusted. Learning concerns are similar to the previous one but they are less critical.
- To invalidate an activity already graded: A learner may not pass an assessment activity and the final score can be highly affected due to the importance of the activity on the assessment. A student could decide to cancel the activity and, therefore, to improve the final score. The cost effort should be less than avoiding the activity since the learner has tried to perform the activity and she has practiced the skills and competences assessed on the activity. On the learning perspective, there are multiple concerns. On the one hand, the student may submit an empty activity

and she can invalidate it with a less effort cost. On the other hand, the learner has tried the activity and she has failed. Therefore, the competences have not been completely acquired.

- To remove constraints in the final mark computation: In some summative models, there are hard constraints that learner must fulfill to pass the course. For example, a grade larger than a certain grade must be obtained from one activity to apply the grading formula. These constraints avoid learners to pass even having high grades in other activities. On the learning perspective, the learner proves with the extra activities knowledge acquisition and if the learner has high grades in some activities, we can assume that most of the competences have been acquired. In this case, the cost effort can be less than avoiding an activity, since this item can be defined as a safeguarding item.
- To see an assessment exercise within an activity previous to the publication date: This item is interesting for activities with a limited time such as final exams. Therefore, the learner could receive the statement of an exercise before the publication date to try to solve it. On the learning perspective, the student practices the competences before the activity and, therefore, the objective to learn is also met. In this case, the cost should be taking into account the points assigned to the exercise in the activity and the cost effort assigned to items related to extra points.
- To change the type of assessment activity: Some courses deploy different activities depending on the level of expertise or the average mark of the learner. Some activity can be more complex such as a final exam, and others can be less complex such as a multiple-choice questionnaire. This item enables the option to change the type of assessment activity on the fly even if the constraints to activate the less complex activity have not met. The cost should be proportional to the type of activity (final exam, continuous assessment activity, etc.). On the learning perspective, the assessment objective of the activity is the same. Therefore, the same knowledge will be assessed.

Note that, this list is not incomplete and more customization items can appear on courses depending on the resources, the activities and the domain of the course. Some of these items may alter the model significantly and it may have a high impact in the grade computation. Therefore, the inclusion of limitations on incompatible items (i.e. items that may generate an invalid model to obtain a score) is a good design decision and it is recommended for all courses.

The evidences are also an important issue of the model. The instructors should decide which ones related to the course can prove student's effort. Here, we summarize a possible list. Similar to the previous list, more evidence may appear depending on the context the model is applied:

- Non-assessment activities: In some courses, there is a list of exercises to practice the contents of the course or some recommended task, i.e. to read some textbook. These activities are not considered in the assessment, since they are recommendations in the formative part of the course. However, the student invests a significant effort doing these activities and this effort can be collected.

- **Additional assessment activities:** Some assessment activities can be defined to practice some contents of the course, i.e. the use of a specific tool. The effort can also be taken into account.
- **Presence on communication channels and discussions:** Some students are active in the communication channels of the course, such as, discussion forums to ask questions and help other learners. This contribution to the collaborative knowledge construction can also be collected.

This is a brief list that can be enlarged with further evidences. Instructors will need instruments to collect them. Automated computer-assisted instruments can be used in online and on-site learning. These tools are preferable since they reduce overheads on manual administration tasks. However, simple instruments such as spreadsheets or manual reports could also be used.

This model can be hard to assimilate for instructors and it can be difficult to explain to students. Instructors are used to define the grading model with some constraints. In some cases, some extra points are rewarded based on the progression of the learner during the course. The proposed model is an extension of the simple reward system of giving points. A student can make an extra effort during the course such as practicing different exercises and the summative model can be customized based on this effort. Another concern in the learning perspective is that student can avoid assessment activities where some critical content of the course is evaluated. Therefore, instructors are unable to assess this knowledge. Assuming a precise effort assignment for each item, if the learner can select an item, this implies that the learner has invested an effort in some parts of the course that should compensate entirely or partially the criticality of the activity. Moreover, similar to other summative models, it is not mandatory to acquire all the competences and learning outcomes of the course. Most of the models, a mark larger than the half of the total points is sufficient to pass the course.

If we refer to students, the model can be described as a reward system based on continuous assessment evaluation. More proactive students will be able to select some benefits compared to learners that only perform the mandatory assessment activities.

4 Personalized Summative Model

After knowing which customization items can be applied and which evidences can be collected, we are ready to define the personalized summative model mathematically.

As aforementioned, the objective of the model is to provide to the learner a customization method to add preferences on his summative model based on the effort exerted during the course. The objective of this section is to introduce the model and define the mathematical constraints.

The Personalized Summative Model (PSM) has two main components:

1. A set of customization items¹ (I) on the assessment. These items are related to all the elements of the assessment and in all parts of the course. For instance, an assessment activity, an exercise of an assessment activity, some extra points in some assessment activity, among others.

¹ Hereafter, we will indistinctively talk about customization items and items.

2. A set of evidences (E) that proves the effort exerted by the student. As we described previously, not all students should have the same customization items. Therefore, the model needs to capture the effort employed by the student regarding evidences. An evidence can be any set of data that prove effort such as the evaluation of an activity, meaningful messages in a discussion board or forum; or traces of the learning material the learner has read, among others.

Based on these components, we can define some constraints onto the model.

Definition 1: Given a set of customization items $I = \{i_1, i_2, \dots, i_n\}$, the compatible customization item matrix M defines the set of compatible customization items that can be selected for a model such that

$$\forall m_{ij} \in M: 1 \leq i, j \leq n \wedge m_{ij} = \begin{cases} 1 & \text{if } i_i \wedge i_j \text{ are compatible} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

The first definition defines the set of customization items and the relationship among them. The matrix defines the set of items that are compatible among them and it helps to avoid the selection of preferences that could generate an invalid model. For instance, for the model of Equation 1, the matrix would set as incompatible the items related to avoid the final project and the final exam.

Definition 2: Given a set of customization items $I = \{i_1, i_2, \dots, i_n\}$ and a compatible customization item matrix M , a learner can select a set of customization items $SI = \{si_1, si_2, \dots, si_k\}$ such that,

$$\forall si_i, si_j \in SI: si_i, si_j \in I \wedge m_{ij} = 1 \wedge k < n \quad (3)$$

A learner should be able only to select items inside the model and only the compatible ones. There is no point to select incompatible items, since the summative model will be inoperative to assess the knowledge acquired in the course.

Definition 3: Given a set of evidences $E = \{e_1, e_2, \dots, e_m\}$ and a set of completed evidences by a learner $SE = \{se_1, se_2, \dots, se_h\}$,

$$\forall se_i \in SE: se_i \in E \wedge h \leq m \quad (4)$$

Let us assume that we can generate evidences of the acquired learning outcomes, skill or competences acquired by the learner. Similar to Definition 2, we specify the relationship between the set of available evidences and the evidences triggered by the student. Note that, these evidences can be specified in multiple ways with respect to the competences of the course, one(evidence)-to-one(competence), many-to-one or one-to-many. This correspondence should be created manually by defining all possible evidences of each course.

Previous definitions define the set of components the learner can collect but it remains to define the relationship between customization items and evidences. Any learner requires a certain amount of effort to acquire knowledge. This effort can be applied on reading a learning material, by performing an assessment activity or by

solving exercises not associated with any assessment of the course and it can be estimated by using the techniques described in Section 2.2.

This effort can be associated with both components: customization item and evidences. In a customization item, it can be seen as the equivalent effort a learner normally invest to achieve the item. In an evidence, it can be viewed as the necessary effort to complete or trigger the evidence.

Note that, the estimation of these efforts is not easy, it may depend on of the course and it also should take into account the type of learner (i.e. a new learner or a repeater do not need the same effort to complete an evidence [20]). This study is left as future work. Here, we assume a manual assignment of effort to customization items and evidences.

Let us assume that we have estimated the effort required to trigger each evidence and the effort associated with each customization item. The following definitions state that each evidence and customization item has an estimation effort coefficient.

Definition 4: Given a set of customization items I , there is a set of effort coefficients $C_I = \{c_1, c_2, \dots, c_n\}$ that establish the effort cost needed to select each customization item such that

$$\forall i_i \in I, \exists c_j \in C_I: c_j > 0 \wedge i = j \quad (5)$$

■

Definition 5: Given a set of evidences E , there is a set of effort coefficients $C_E = \{c_1, c_2, \dots, c_m\}$ that establish the effort cost needed to complete each evidence such that

$$\forall e_i \in E, \exists c_j \in C_E: c_j > 0 \wedge i = j \quad (6)$$

■

Note that, Definition 4 and Definition 5 can be applied to any set of customization items and evidences. That is, they can also specify the set of effort coefficients for the selected item and triggered evidences for a learner. The effort coefficients for the selected customization items and completed evidences by a learner are denoted as C_{SI} and C_{SE} respectively.

Property 1: Given a set of customization items I and a set of evidences E for a PSM model, a set of selected customization items SI and a set of triggered evidences SE for a learner; we can define the next constraint for the respective sets of effort coefficients C_I , C_E , C_{SI} , and C_{SE} :

$$\sum_{c_i \in C_{SI}} c_i \leq \sum_{c_i \in C_{SE}} c_i \leq \sum_{c_i \in C_E} c_i < \sum_{c_i \in C_I} c_i \quad (7)$$

■

Property 1 specifies that the total effort associated with the selection of customization items is less or equal to the total effort of the triggered evidences and, consequently, to the total effort required to complete all evidences. Additionally, the property also specifies that the total effort cost of all items of the model is always larger

than the total effort of the selected ones. This last constraint ($\sum_{c_i \in C_E} c_i < \sum_{c_i \in C_I} c_i$) is crucial for the model since we clearly define that the model does not allow to select all items. In other words, the model will have some incompatible items that cannot be selected at once by a learner.

Property 2: Given a set of customization items I and a set of evidences E for a PSM model with the corresponding sets of effort coefficients C_I and C_E , we define the complexity coefficient $\alpha \in [0 \dots 1]$ such that

$$\sum_{c_i \in C_E} c_i < \alpha \sum_{c_i \in C_I} c_i \text{ and } \forall c_i \in C_I: c_i \leq \sum_{c_j \in C_E} c_j \quad (8)$$

■

The complexity coefficient α constraints the number of items that can be selected from the model based on the total effort of all items. On values near to 1, mostly all items can be selected (all items cannot be selected as stated by Property 1), meanwhile for values near to 0, a small set can be selected. The second constraint sets a lower bound for the total effort of evidences requiring that all items should be selectable. That is, there is no item with an effort larger than the total effort of evidences.

Note that, this is a simple model that establishes a unique effort estimation through the course. This implies that the acquisition of different knowledge is abstracted to a unique effort value and this effort can be exchanged for any item (i.e. even items without any relationship with the evidences which provided the effort). However, this model can be extended to various effort estimations allowing a more complex model. This second approach could refine the correlation between different effort estimation and items.

Also, we can enhance the model by defining different coefficient efforts to each learning outcome or competence. Different coefficient efforts can provide a more fine-tuned model where evidences will be associated only with the set of customization items related to the same learning outcome. However, we consider that these more complex approaches at this stage of the research are out of the scope.

5 Data Structure Information

This section describes the data structure for storing the information proposed for the personalized summative model. Note that the design has been performed from scratch based on the necessary knowledge to be stored. The proposed structure contains the minimal information required to collect all the completed evidences and select the desirable customization items.

A UML representation can be seen in Figure 1. The model has been designed similar to a payment transaction system (i.e. banking system) where the currency is the learner effort. The main entity is *transaction* that can be seen as the effort account of the learner. It is the responsible to store all the customization items and evidences that a learner has selected and completed respectively. The entity stores each transaction the learner performs and it keeps the balance of the available effort. Note that, the

available effort is restricted to be always positive. The model facilitates the tracking of the effort by assigning a positive effort to evidences and a negative effort to the customization items. Thus, a positive effort on transactions denotes effort acquired by completing evidences and not yet consumed by selecting customization items.

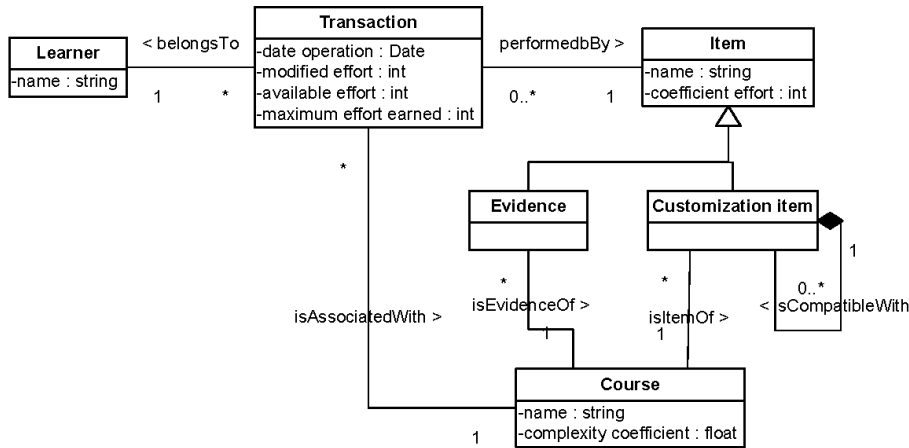


Fig. 1. Data structure of the Personalized Summative Model

A transaction *belongs to* a learner and it *is associated with* a course. Thus, a learner can have different active PSM models but each one is associated with a unique course. Note that, a transaction can be only *performed by* two types of *items*: *customization items* and *evidences* and they have an associated coefficient effort as we defined in the previous section. Note that, the model also stores the compatible set for each course using the *is compatible with* relation.

Finally, each evidence *is evidence of* a course and a customization item *is item of* a course. These two relations help to specify the set of evidences and items for each course. Here, we assume that evidences and items are different for each course and this assumption simplifies the model significantly when the effort is assigned to each item. Although evidences and customization items can be similar across courses, if we assume that they are different, the effort assignment is set to each item individually. The *course* entity also defines the complexity coefficient to limit the maximum effort that can be achieved by triggering evidences and the transaction entity keeps the maximum effort earned by each learner to facilitate the checking of the constraint.

Currently, the model does not store relationship with learning outcomes. We assume that the manual definition of evidences subsumes the set of learning outcomes of the course implicitly.

6 Analysis Opinion Survey

The PSM model has not been yet implemented since an analysis of the customization items and evidences that can be applied should be performed prior to their utiliza-

tion. Similarly, the effort estimation must be further analyzed. Moreover, depending on the context of the course (online or on-site), a system should be designed and deployed to facilitate the utilization of the model.

At this stage, we have performed a qualitative evaluation of instructors’ and learners’ opinion as potential users of the PSM model. Precisely, different surveys have been used. Instructors are queried about the potential of the model and learners are questioned based on the set of items we have defined in the previous section.

6.1 Survey to instructors

For this purpose, a survey was performed to give an opinion of the new model. The Likert scale was used ranging from 1 to 5, being the larger value five meaning that they strongly agree with the statement. At the end of the survey, there is also an open question to gather critical comments and valuable opinions not collected in the previous statements.

The statements were the following ones:

1. The model is conceptually easy to understand.
2. The model can be used in a higher education context.
3. The model can be easily deployed on a course.
4. The proposed customization items are easy to understand.
5. The proposed customization items are acceptable based on the learning perspective.

The survey has been performed on ten teachers of courses related to high education courses in Computer Science. We are aware that it is a too small sample and a large-scale survey should be done to obtain conclusions. The objective of this short survey was to get first impressions of the new model and detect advantages and drawbacks to improve it. An extensive survey is planned as future work. Results are summarized in Table 1.

Table 1. Results of the instructor's survey (1 to 5 scale, 5 is better)

Statement	Instructor										Avg.
	1	2	3	4	5	6	7	8	9	10	
1	4	5	4	3	5	4	4	5	4	5	4,3
2	5	5	5	4	5	5	5	5	4	5	4,8
3	3	4	3	2	4	3	4	4	1	3	3,1
4	5	4	5	3	5	5	5	5	3	5	4,5
5	4	5	5	3	4	4	4	5	3	4	4,1

The first statement evaluates the complexity of the model based on the objective of the model and the mathematical constraints. Instructors agree that is easy to understand.

The second statement evaluates the perception whether this model can be applied in a high education context. All instructors agree with the statement. On the open

question an instructor argues that it should be interesting to describe how the model can be combined with a formative model to boost the instructional process.

The third statement obtains the lowest rating. At this stage the instructors do not see clearly how they can apply this new model to a course. In the open question, they argue different concerns: 1) it is hard to define the customization items and the evidences. 2) there is no instrument to estimate the effort cost of the customization items and the evidences. 3) The effort is subjective and it can be overestimated or underestimated for different items and evidences and for each learner.

The fourth statement evaluates the complexity of the defined customization items. The instructors mostly agree with the statement that they are easy to understand. However, in the open question, there is a concern related to how some of them, such as cancel an activity or an exercise, can really be managed during the course. Also, an instructor found a pitfall in the customization item “to cancel constraints”: A learner can submit an empty exam equivalent to avoid the exam with less effort.

Finally, instructors partially agree that the customization items can be acceptable based on the learning perspective. There is a concern in the open question related to which learning implication has that a learner can avoid or cancel an activity. Is it acceptable? or has the learner to prove (be assessed) the knowledge in another way?

At this stage, we consider that this appraisal is acceptable. Instructors accept the model as a possible one. However, they have several concerns that should be discussed in the next steps of the research. A previous analysis should be performed on courses since the customization items and the evidences have to be defined. Although, the list described in the previous section cannot be applied straightforwardly, each course has particularities that imply that the list of available items and evidences should be specifically defined.

Additionally, instructors need more instruments. In particular, they need specifically how to estimate the effort cost, how to manage the effort performed by each learner and how to manage the selection of evidences for each student. These instruments should be deployed by a specific system which should support the complete model.

6.2 Survey to students

Another questionnaire has been performed on students. The aim of the study is to query which customization items will be preferred by the students since some of them may not be attractive to them.

The survey has been performed on the course Computer Fundamentals of the Degree of Computer Science of the Open University of Catalonia that is a first-year course with a large number of students (around 350 students per semester). The contents of the course are related to the introduction of digital systems design.

The course is composed of three continuous assessment activities (CAA), one final project (FP) at the end of the semester and an on-site final exam (FE). The CAA are related to contents of numeral systems, combinational circuits and sequential circuits. The final project is intended for practicing all the knowledge acquired during the course to build a sequential circuit based on a finite state machine. Finally, the final exam assesses particular knowledge acquired during the CAA.

Table 2. Results of the student's survey (0 to 5 scale, 5 is better)

Average score for activity / item	CAA1	CAA2	CAA3	AC	FP	FE
To add extra points (1 point)	1,3	2,1	3,0	3,6	3,9	4,3
To extend submission date (2 days)	1,5	2,2	2,8	-	3,6	-
To avoid an activity	2,0	3,3	3,5	0	4,4	4,7
To avoid 1 exercise within an activity	2,0	3,1	3,3	-	4,1	4,6
To cancel an activity	2,1	2,5	3,0	0	3,9	4,5
To nullify constraints	-	-	-	-	3,8	4,1
To see exercise statement previous publication date (1 day)	0	0	0	-	0	4,0
To change type activity (problem-based exam to multi-choice questionnaire)	-	-	-	-	-	4,2

The final mark is computed based on Equation 1 where we can observe that the continuous assessment (CA) is optional. However, the FP and FE are mandatory with a minimum score of 4 points.

Here, the customization items defined in Section 3 have been distributed on each activity. Note that, some items have been avoided since they were not applicable in the activity, i.e. extension of the submission date is not applicable on an on-site final exam since the exams are institutionally scheduled and this schedule cannot be modified. A Likert scale ranging from 0 to 5 has also been used, being the larger value 5 meaning that they strongly agree with the relevancy of the item; and 0 meaning that the item is not relevant for the activity. For each item, the average score on the Likert scale has been computed.

The survey has been introduced as an additional non-mandatory activity (not included in the assessment of the course) at the end of the semester (near the submission of the final project). Note that, at this stage of the course, learners know how the course is distributed and the difficulty inherent in each activity (except the final exam) and the final mark computation. Therefore, this experience is excellent to analyze where customization items could be better deployed. The survey was answered by 143 out of 345 students (41%).

As we can observe in Table 2, learners give more importance to items on activities that have more impact in the final mark computation, i.e. CA, FP and FE. This is a consistent result. The objective of the item is to customize the summative model. Therefore, students prefer to customize more complex activities that have more impact on final mark computation. The more relevant item is to avoid the final exam followed by avoiding one exercise of the final exam. To cancel the exam and to avoid the final project are the following preferences.

Another interesting result is that some items are not relevant at all. To see an exercise statement of a CA activity previous to the publication date or to cancel or not to perform a CA activity are not selected by any student. Note that, not to perform the CA is inherent on the grade computation formula described by Equation 1.

This survey gives first insights that these items are relevant for students. However, we are aware that these results may differ from other courses and different final mark computation models.

7 Conclusions and Future Work

In this paper, we have presented a novel summative model based on learner preferences. Learners are able to customize their summative model based on their level of engagement and effort during the course.

Mathematical constraints to build the model and the learning perspective of the model have been described. Additionally, a survey to instructors and learners has been presented to analyze first opinions about the new model and the defined customization items.

Currently, the model has many unsolved challenges. First, the effort should be estimated. The models presented in Section 2.2 can be used or even a combination of them. A statistical model will be constructed as an instrument to estimate the effort for each course efficiently. Evidences should be carefully defined since they may have different formats or unstructured and their collection should be analyzed. Online and on-site learning should also be taken into account when defining them. Online learning will be easier to collect and process evidences since VLE combined with learning analytics systems can provide deep reports for each learner. Whereas on-site learning can be more difficult since evidences should be gathered manually. However, it is feasible to gather effort in face-to-face settings as stated in the literature described in Section 2.2.

More complex challenges appear when learning outcome are assessed. The current model assumes a unique effort estimation for all learning outcomes and, therefore, the effort collected on evidences could be invested in any customization item. But evidences should be classified based on the type of learning outcome where they have been collected and the effort received should be consumed on preferences compatible with the outcome. Thus, the effort has a certain *label* associated based on the practiced knowledge performed during the completion of the evidence.

Finally, this model may have some criticism from academic institutions or quality agencies since it clearly states that some competences it may not be learned or assessed. However, as we described in the introduction, this model is not far from current models where not all the competences are assessed in the activities (i.e. the instructor decides which knowledge is really assessed on each activity) and where the learner usually needs the half of the total score to pass the course. The difference is that in the PSM model is the learner who decides which knowledge is avoided during the assessment activity. All these open challenges will be addressed as future work.

8 Acknowledgment

This work was funded by the Spanish Government through the project: TIN2013-45303-P "ICT-FLAG: Enhancing ICT education through Formative assessment, Learning Analytics and Gamification".

9 References

- [1] Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35(4), 435-448. <https://doi.org/10.1080/02602930902862859>
- [2] Liu, Z., & Cercone, N. (2016). Design of an adaptive learning system and educational data mining. *Data Mining and Learning Analytics: Applications in Educational Research*, 81.
- [3] Wolf, C. (2002). iWeaver: towards an interactive web-based adaptive learning environment to address individual learning styles. *European Journal of Open, Distance and E-Learning*, 5(2).
- [4] Zhang, J. C. (2016). Adaptive Learning Environment System Based on Multi-event Driven Technology. *International Journal of Emerging Technologies in Learning*, 11(11). <https://doi.org/10.3991/ijet.v11i11.6250>
- [5] Knight, P. T. (2002). Summative assessment in higher education: practices in disarray. *Studies in higher Education*, 27(3), 275-286. <https://doi.org/10.1080/03075070220000662>
- [6] Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education*, 31(2), 199-218. <https://doi.org/10.1080/03075070600572090>
- [7] Passerini, K. & Granger, M. J. (2000). A developmental model for distance learning using the Internet. *Computers & Education*. 34(1), 1-15. [https://doi.org/10.1016/S0360-1315\(99\)00024-X](https://doi.org/10.1016/S0360-1315(99)00024-X)
- [8] Roberts. T. S. (2006). The use of multiple choice tests for formative and summative assessment. In *Proc. of the Australasian Conference on Computing Education*, pp 175-180.
- [9] Bennett, R. E. (2010). Cognitively based assessment of, for, and as learning (CBAL): A preliminary theory of action for summative and formative assessment. *Measurement*, 8(2-3), 70-91. <https://doi.org/10.1080/15366367.2010.508686>
- [10] Hernández-Leo, D. & Moreno Oliver, V. (2014). Fine-tuning formative and summative assessment in Bachelors' Final Projects. *2014 International Symposium on Computers in Education (SIIE)*, pp. 41- 46. <https://doi.org/10.1109/SIIE.2014.7017702>
- [11] Sorva, J. & Sirkiä, T. (2015). Embedded questions in e-books on programming: useful for a) summative assessment, b) formative assessment, or c) something else? In *Proceedings of the 15th Koli Calling Conference on Computing Education Research*, pp. 152-156. <https://doi.org/10.1145/2828959.2828961>
- [12] Zolotaryova, I & Plokha, O. (2016). Serious games: Evaluation of the learning outcomes. *2016 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET)*, pp. 858-862. <https://doi.org/10.1109/tcset.2016.7452206>
- [13] Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, 4(3), 365-379. <https://doi.org/10.1080/0969594970040304>
- [14] Taras, M. (2005). Assessment—summative and formative—some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466-478. <https://doi.org/10.1111/j.1467-8527.2005.00307.x>
- [15] Kommers, P., Grabinger, S. & Dunlap, J. (1996). Hypermedia Learning Environments. *Instructional Design and Integration*. Lawrence Erlbaum Associates.
- [16] Pellegrino, J., Chudowsky, N. & Glaser, R. (2001). Knowing what students know: The Science and Design of Educational Assessment. *National Academy of Sciences*, National Academy Press, 2001.

- [17] Brown, R. W. (2004). Undergraduate summative assessment experiences, *34th Annual Frontiers in Education, FIE 2004*, (2), F3G-5-9.
- [18] Navrat, P. & Tvarozek, J. (2014). Online programming exercises for summative assessment in university courses. In *Proceedings of the 15th International Conference on Computer Systems and Technologies (CompSysTech '14)*, pp. 341-348.
- [19] Alotaiby F. T. & Chen, J. X. (2005). Generic summative assessment functional model. *Sixth International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing and First ACIS International Workshop on Self-Assembling Wireless Network*, pp. 340-345. <https://doi.org/10.1109/snnpd-sawn.2005.39>
- [20] Swinton, O. H. (2010). The effect of effort grading on learning, *Economics of Education Review*, 29(6): 1176-1182. <https://doi.org/10.1016/j.econedurev.2010.06.014>
- [21] Swinton, O. H. (2015). An A for Effort. *The American Economic Review* 105(5): 616-620. <https://doi.org/10.1257/aer.p20151116>
- [22] Biddle, S., & Goudas, M. (1997). Effort is virtuous: Teacher preferences of pupil effort, ability and grading in physical education. *Educational Research*, 39(3), 350-355. <https://doi.org/10.1080/0013188970390310>
- [23] Wolters, C. A. (1999). The relation between high school students' motivational regulation and their use of learning strategies, effort, and classroom performance. *Learning and individual differences*, 11(3): 281-299. [https://doi.org/10.1016/S1041-6080\(99\)80004-1](https://doi.org/10.1016/S1041-6080(99)80004-1)
- [24] Schuman, H., Walsh, W., Olson, C., & Etheridge, B. (1985). Effort and reward: The assumption that college grades are affected by quantity of study. *Social Forces*, (4), 945-966. <https://doi.org/10.2307/2578600>
- [25] Hill, L. (1991). Effort and reward in college: A replication of some puzzling findings. In W. N. James (Ed.), *Replication research in the social sciences*. pp. 139-149. Sage Publications, Inc.
- [26] Michaels, J. W., & Miethe, T. D. (1989). Academic effort and college grades. *Social Forces*, (1), 309-319. <https://doi.org/10.2307/2579230>
- [27] Dobkin, C., Gil, R., & Marion, J. (2010). Skipping class in college and exam performance: Evidence from a regression discontinuity classroom experiment. *Economics of Education Review*, (4), 566-575. <https://doi.org/10.1016/j.econedurev.2009.09.004>
- [28] Kanfer, R., Ackerman, P. L., Murtha, T. C., Dugdale, B., & Nelson, L. (1994). Goal setting, conditions of practice, and task performance: A resource allocation perspective. *Journal of Applied Psychology*, 79(6), 826. <https://doi.org/10.1037/0021-9010.79.6.826>
- [29] Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, 52, 139-183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [30] Paas, F. G., & Van Merriënboer, J. J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of educational psychology*, 86(1), 122. <https://doi.org/10.1037/0022-0663.86.1.122>
- [31] Fisher, S. L., & Ford, J. K. (1998). Differential effects of learner effort and goal orientation on two learning outcomes. *Personnel Psychology*, 51(2), 397-420. <https://doi.org/10.1111/j.1744-6570.1998.tb00731.x>
- [32] Yeo, G. B., & Neal, A. (2004). A multilevel analysis of effort, practice, and performance: effects of ability, conscientiousness, and goal orientation. *Journal of Applied Psychology*, 89(2), 231. <https://doi.org/10.1037/0021-9010.89.2.231>
- [33] Baneres, D.; Baró, X; Guerrero, A. E.; Rodriguez, M. E. (2016). Adaptive e-Assessment System: A General Approach. *International Journal of Emerging Technologies in Learning (iJET)*, 11(7):16-23. <https://doi.org/10.3991/ijet.v11i07.5888>

10 Author

David Baneres is with IT, Multimedia and Telecommunication Department at the Open University of Catalonia, Rambla del Poblenou 156, Barcelona 08018, Spain. His research interest involves e-learning innovative methodologies and intelligent tutoring systems. He is also a member of the SOM Research group at UOC (*som-research.uoc.edu*).

Article submitted 17 May 2017. Published as resubmitted by the author 20 June 2017.