

Article

# Evaluation for Teachers and Students in Higher Education

Lineth Alain Botaccio <sup>1</sup>, José Luis Gallego Ortega <sup>2</sup>, Antonia Navarro Rincón <sup>3</sup> and Antonio Rodríguez Fuentes <sup>4,\*</sup>

<sup>1</sup> Department of System Information, Faculty of Systems Engineering, The Technological University of Panama, Panama City 0819-07289, Panama; lineth.alain@utp.ac.pa

<sup>2</sup> Department of Didactics and School Organization, Faculty of Educational Sciences, The University of Granada, 18071 Granada, Spain; jlgalleg@ugr.es

<sup>3</sup> Department of Didactics of Language and Literature, Faculty of Education and Sport Sciences of Melilla, The University of Granada, 52005 Melilla, Spain; anavarro@ugr.es

<sup>4</sup> Department of Didactics and School Organization, The University of Granada, 18071 Granada, Spain

\* Correspondence: arfuentes@ugr.es

Received: 1 April 2020; Accepted: 1 May 2020; Published: 15 May 2020



**Abstract:** It is time to undertake changes in the evaluation methods we use, especially in higher education. These changes in the actors responsible for evaluation would combine hegemonic traditional evaluating processes with other, more democratic modalities, which would turn the predominantly institutional rating purposes of evaluation into a learning experience, and develop a competence in evaluation in students. Only in this way can coherence be achieved within the context of the student's initiative and the construction of their learning, mainly because of their real empowerment in the didactic process, either individually or in groups. A virtual platform has been developed to avoid increasing the teaching load and it is exposed in this work. The platform has been built and validated by potential users following the design-based research model. Its description, as well as its results, are explained. Regarding the description, two interfaces are mentioned—one for teachers and another for students. Concerning its validation, the results of this quantitative and qualitative study confirm its functionality as a valid tool for evaluation. It is predicted that the utilization and impact of this tool will not only be beneficial for the evaluation dimension, but also for the overall improvement of the teaching experience.

**Keywords:** higher education; evaluation; evaluators; educational software interface

## 1. Introduction

It is necessary to reconsider the currently used evaluation methods [1], due to their impact on and transcendence in the enhancement of the educational reality, and didactic process agents, i.e., teachers and students. If the aim of current didactic trends is to encourage students to play a more active role in the construction of learning, this should be endorsed through the empowerment thereof in the evaluation process. The same applies to metacognitive knowledge of learning or meta-learning [2].

This requires overcoming traditional evaluation models by taking them away from the hands of the teacher and encouraging the involvement and responsibility of the students, and even of other agents in such an undertaking [3–6]. This has been achieved by providing these aforementioned agents, including students, with such an opportunity [7–9], which has had favorable results in the form of innovative experiences [10–13].

In the previous line of reasoning, other participatory and democratic evaluation typologies were recognized, which, far from being inferior or exclusive, are pertinent and complementary. These

are beneficial due to their contribution to the stimulation of a critical and constructive attitude. Furthermore, such typologies are beneficial by attracting full and conscious attention to learning (promoting meta-learning and future learning by improving learning capacity) and beyond (stimulating their evaluative capacity for their later professional development in teaching) from students.

The preceding arguments have impacted contemporary university education, where students are responsible for building their own learning. If an active initiative is in place for teachers and students, they have to be empowered through their own evaluation [14]. Self-evaluation, in the conceptual framework of evaluation being seen as a motivating factor for improvement [15,16], should be perceived as an impulse toward a self-critical attitude and reflection that contributes to personal and professional maturity. If evaluation requires some maturity, the university environment should be the most propitious place for its application [7], and even more so if those being evaluated are teachers in training whose evaluation competence is still developing [10].

In addition, if collaborative teamwork is demanded within homogeneous evaluation modalities, a shift towards the empowerment of the class group and operational working groups is required, granting them levels of evaluating responsibility. This peer-to-peer evaluation (in pairs or groups) and co-assessment gain ground in the university context after the implementation of active methodologies derived from the European higher education area (EHEA) [4].

The benefits of the aforementioned practices include greater involvement, responsibility, communication, and a critical attitude by the students not only when the time comes to evaluate their teacher but also while they are being taught [3,17–19].

Furthermore, the multi-criteria evaluation proposal is not limited to the evaluation agents but includes more widely accepted elements, such as:

- (a) At different points in time—not only at the end but also at the beginning of and during the process;
- (b) With different tools, complementary to the anachronistic monopoly of the traditional exam [20];
- (c) At different evaluated dimensions, commonly only tangible knowledge, to which less obvious procedures and attitudes have been added, as well as other more complex competencies [21,22];
- (d) Finally, with different purposes—not only to reward, sanction, categorize or school the student (modality, promotion, etc.), but also to redirect efforts and raise awareness.

It has been argued that “evaluating is not just to qualify but to verify day by day that the teaching approach is bringing about the desired effect and the learning is blossoming. This implies that the teaching of the teacher is contributing to the development of the student’s learning” [10] (p. 351). In addition, evaluating is to bring an improvement in their performance [23], as well as self-regulation of knowledge and readjustment of teaching efforts: “feedback” and “feedforward” [24] (p. 45) or “feedback” and pro-feeding [25] (p. 2), with teachers also reporting self-regulation and learning for themselves as a professional challenge.

However, there is a discordance between the previous proposals and the current evaluation practices. In the latter, a perpetuation of the traditional hetero-evaluation methodology is observed [1]. Its starting and arrival point is evaluation of the students by the teacher, in a unidirectional journey conducted, predominantly, by traditional examinations [20] to gauge the reception of the studying process, instead of the learning process itself [26].

Given this disagreement, it is mandatory to broaden the paths leading to evaluation since, far from being incompatible or exclusive, they can coexist and combine perfectly to enrich the process, not only as evaluators but also as a global teaching–learning experience. Teachers’ reluctance against a combination of evaluations can be justified for several reasons:

- (a) Persistent rigidity of educational institutions and systems;
- (b) Traditional attitudes toward teacher training;
- (c) The additional complexity of recording diverse scores and their weighted calculation to obtain the overall score, as well as the lack of resources to facilitate this.

Based on this last point, this work has dual objectives. Firstly, presenting the PLEVALUA (evaluation platforms) virtual platform (outcome of a teaching innovation project called "PLEVALUA: Combined Assessment Platform: hetero-evaluations, self-evaluations and co-evaluations of university students") so that all agents, teachers and students, individually or in groups, can introduce scores of developed tasks, which would be useful to reduce the complexity of registration and systematization of evaluations of different activities, moments and evaluation agents to obtain a single numerical mark for every student and/or group.

Additionally, it is encouraged to check the functionality of the platform for its multiple evaluation nature with a sample of teachers and university students, duly instructed for its use. This test will study the perceptions of the sample group on these multiple evaluations as well as its facilitation through the platform.

As a hypothesis, one can venture that the participants in the research are going to welcome the multi-evaluation proposal as well as the facilitating platform, whilst the teachers are conscious of the advisability of combining different marks and the use of the platforms, and that the students are also digital natives and their familiarization and attraction to the ICT (information and communications technology) is a fact.

The hypothesis is based on the current recognition and use of digital platforms in university education (cf. revision of [27]). Some platforms allow for the collection of works for their evaluation by the educator, however, it is true that the use of these platforms for the evaluation of learning by different agents (teachers and students) is unusual [25]. It draws from the recognition and the current employment of digital platforms in higher education (cf. revision of [27]). Some platforms allow collecting tasks for evaluation by the teacher. However, the use of these tools for the evaluation of learning by different agents (teachers and students) is rather unusual. It has not been possible to locate studies such as this concerning the proposal, description of the platform, and its validation through the design of the research detailed below.

## 2. Material and Methods

This proposal is part of the methodology of the design-based research (DBR), derived from the research–action approach from the field of engineering and other applied sciences [28]. It consists of the creation of an online platform to propitiate plural evaluation, taking into account the university context and the current demands for evaluation in higher education.

On one hand, the development of the platform PLEVALUA (University of Granada, Spain, and the Technological University of Panama, Republic of Panama, code number 1906241268126, PLEVALUA "Plataforma para la evaluación múltiple universitaria: realizada por el heteroevaluaciones, coevaluaciones y autoevaluaciones del alumnado universitario", Spain) has been conducted according to the methodology of software design proposed and validated by Roger Pressman [29], which proposes three phases:

1. A definition process made up of a sub-phase of requirements of potential users and planning of activities and times.
2. A development process, in which two sub-phases are addressed:
  - The design of a pattern, developed using the corresponding programming language.
  - Software maintenance to optimize the product in its non-final version.
3. A constant maintenance process, where technical problems of the final version will be solved, and, where appropriate, it is replaced by an upgraded version, in a regular cycle.

Two techniques were used following the DBR design for the validation of the functionality of the platform, which generated an effective mixed methodology, recognised in the field of educational research:

- Content analysis, based on voluntary and anonymous statements issued by participants, teachers, and students, duly instructed and experienced in the use of the platform.
- Statistical analysis based on the data obtained using a self-filled Likert multiple response estimation scale.

### 2.1. Participants

It is not possible to obtain a representative and random sample but instead to obtain a convenience sample from research participants, including teachers (TE) and students (ST):

- 30 teachers from the University of Granada from different specialties attended the specific course on this topic (entitled “Combined evaluation of students, classmates, and teachers through PLEVALUA digital platform”, organized by the Quality, Innovation and Prospective Unit of the University of Granada (2018)). They each had under 12 years of teaching experience ( $M = 6.50$ ,  $sd = 2.76$ ) and there were more women (56.67%) than men (43.33%).
- Regarding the students, a total of 140 students working towards a primary education teaching degree from the same university, at their second (41.43%) and fourth stage (58.57%) took part, which implies that they already had some university experience ( $M = 3.17$ ,  $sd = 1.72$ ). The proportion of women compared to that of men is even greater in this case (74.62% and 25.38%, respectively), which correlates with the reality of the classrooms in these studies.

All participants declared knowledge and familiarity with digital platforms. In sum, they amounted to 170 participants: 17.65% of the teaching group and 82.35% of the student group. These are the participants who have expressed their opinions on the multi-evaluator modality and the platform, and they submitted to quantitative analyses.

Similarly, everyone had the opportunity to reflect their statements through the suggestion box on the platform, after ensuring the anonymity of the scales and the platform, despite recording their data, through the process of the anonymous dump. The reality was, however, that only 45 of them added any valid statement to the mailbox, after deleting two student reviews due to their lack of clarity of interpretation (one that only indicated “yes” and another that only indicated “ok”). Of the total, 26.67% are teachers and 73.33% are students, of which gender, specialties, and courses are unknown, given the anonymous nature of the research, which is not inconvenient as they will be considered as a single case for this qualitative analysis.

### 2.2. Instruments and Procedure

The suggestion box on the platform was used to evaluate the opinion on the required tasks with regard to perception, access, and use. It is a data collection technique that is part of the model of qualitative research to give voice to students or research participants, in a free and anonymous way. Anonymous because of the subsequent dump of data (declarations); the personal details of the participants in the platform were not incorporated. Suggestions were free since there was not any question or response guideline included, just the encouragement of teachers to express what they wish and only if they wish, as well as the ability to use their own format of suggestions.

Additionally, a Likert-type estimation scale was employed to know the opinion of the teaching staff on the evaluation modality and the benefit of the platform. This scale included four response options (1 = strongly disagree, 2 = disagree, 3 = agree and 4 = absolutely agree). It combines two blocks of paired questions: one of them comprises questions about the perception of evaluation and the other about the possibility of making an evaluation through the platform, which will first be analysed separately and then jointly.

The scale (Table A1) was firstly validated by expert judgment, by the teachers who developed the innovation project in which the work is framed. The scale was validated statistically afterward using Cronbach’s alpha, for the teaching staff:  $\alpha_{total} = 0.75$ , total and by blocks  $\alpha_{multievaluation} = 0.71$

and  $\alpha_{\text{functionality-platform}} = 0.77$ ; and for the group of students:  $\alpha_{\text{total}} = 0.70$ , total and by blocks  $\alpha_{\text{multievaluation}} = 0.68$  and  $\alpha_{\text{functionality-platform}} = 0.73$ .

As for the procedure, the application of PLEVALUA by the two groups has been sought: a) for teachers on the occasion of a training course on evaluation and the platform PLEVALUA and b) for students for the mere use of the platform experimentally in their assessments.

Next, a series of supervised activities with PLEVALUA were requested so that the subjects know its purpose and functionality, the last of them the voluntary expression of opinion through the mailbox on the platform.

In the face-to-face delivery of activities, the scale was implemented on a Google document, one for the teachers and one for the students, with different academic and professional data, but with the same estimation criteria, and both voluntary and anonymous nature. It was highlighted that the interest of the valuation is to validate the use of the platform and its exclusive dissemination in institutional and scientific channels as well as its anonymity.

### 2.3. Analysis of Data

Content analysis was applied for the analysis of student statements taken from the suggestion box. More specifically, information was reduced through the establishment of categories with hindsight, lacking pre-established (deductive) categories that conform to an accepted theoretical model for the evaluation of the use of evaluation platforms. The Nvivo version 12 programme was employed to enable these analyses. Finally, after the triangulation and consensus of categories, four metacategories were established, two for the positives and two for the negatives:

- Multiple evaluation assessment (MEA), which encompasses partial assessments by modalities provided they are conceived as part of the whole.
- Assessment of the use of PLEVALUA (AUP), with all categories on concrete and global aspects of the platform.
- Critical review of the multiple evaluation (CME) in general or of any of its constituent modalities as well as a commitment to some in exclusivity.
- Critical review of the use of PLEVALUA (CUP), difficulties, limitations, lack of functionality, etc. that impact on the task for which it was devised.

For the analysis of the data provided on the scales, descriptive statistical analysis, of central tendency—mean (M) and mode (Mo)—and dispersion (standard deviation—sd), and inferential analysis (Student's t, ANOVA, and coefficients of parametric and non-parametric correlation, depending on the case) were applied, under the IBM SPSS programme version 22, and assuming an error level of 5% ( $p < 0.05$ ). The former responds to the purpose of drawing a profile of the participants on the combined evaluation and use of PLEVALUA. The inferences claim:

- Relating (through Pearson's "r" parametric test to the group of students ( $n = 140$ ) whose data distribution was normal, according to the Kolmogorov-Smirnov (KS) test calculation, and similar variances, according to the homoskedasticity test performed with the Levene "L" test, and in the case of teachers, the non-parametric coefficient of Spearman's rho "ρ" for data ( $n = 30$ ) that despite following a normal distribution, according to KS, does not own the necessary homoscedasticity for the calculation of parametric tests).
- Differentiating (Student's t for the same sample) dependent variables on the perception of the specific evaluation modality and the functionality of the platform for such evaluation modality (two blocks) of the scale.

Additionally, distinguishing positions by groups according to the independent variables (for dichotomous, such as gender and contact with ICT, Student's t for different samples and numerical samples, such as age, as well as polytomous items, as experience, the ANOVA of a factor). With special emphasis on the differentiation between teacher and student group responses, through Student's t for

different samples. The tasks to be evaluated consist of solving practical case scenarios proposed for each topic by the teacher. Since these are presented in class, the students have the opportunity to issue their evaluation. The platform PLEVALUA guarantees anonymity, in such a way that the teacher can see on their profile the set and calculation of multiple scores.

### 3. Results

In the first place, the elaborated platform, its description, and location are presented and then validated by its potential users.

#### 3.1. Platform Description

The platform to compile different evaluations of agents (students of themselves, peers and teachers) and at different times (repeated, continuous or procedural), called PLEVALUA, was created and perfected by Roger Pressman's software engineering methodology (risk agile) [29], as free software located in open mode after registration on the site [http://www.linyadoo.com/plevalua\\_ug/](http://www.linyadoo.com/plevalua_ug/).

Its creation responds to the claimed need for a technological resource for the procedural and combined evaluation of all the agents taking part in the didactic act. In fact, it consists of two interfaces easily differentiated by their role at the time of registration ("user type" in Figure 1):

The image shows a web form titled "Agregar Usuarios". At the top, there is a section for "Importar Usuarios" with a "Seleccionar archivo" button and a text box showing "Ningún archivo seleccionado". Below this is an "Importar" button and another "Ningún archivo seleccionado" text box. The "tipo de Usuario" section has two radio buttons: "Profesor" (selected) and "Estudiante". Below these are several text input fields: "Nombre", "DNI", "Correo", "Teléfono", "Nick", and "Clave de Ingreso". At the bottom of the form, there is a "Guardar" button and a red link that says "VOLVER AL LOGIN".

**Figure 1.** User registration interface (teachers and students). Source: taken from PLEVALUA [30].

The interface for teaching staff allows the creation of a course by adding students either from an Excel list, with the option "select file" in Figure 1, or manually, as seen in Figure 2. The command menu is on the left side (Figure 2). This includes the following commands: "group practices" to generate practices and groups; "evaluation" for quantitative and qualitative assessments; "see evaluations" to visualize the scores and observations; "update" to load new practices and "exit" to log out with a logged-in user.



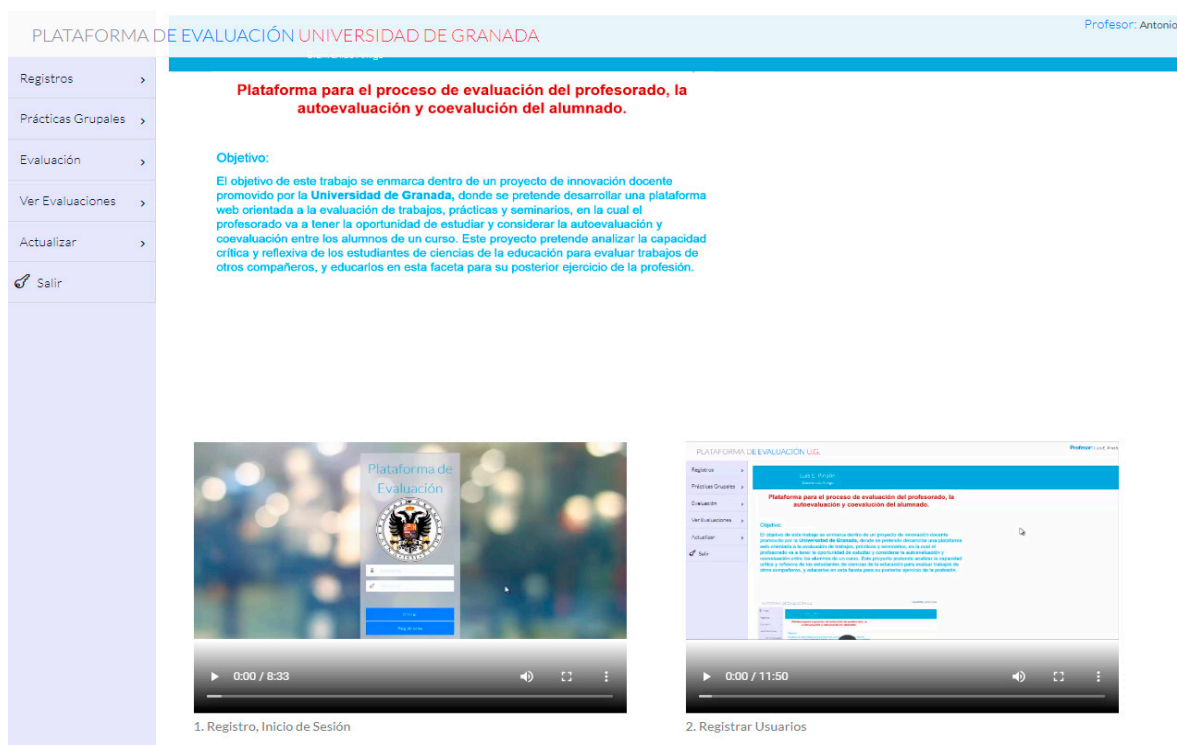


Figure 2. Start of PLEVALUA in the role of teacher. Source: taken from PLEVALUA [30].

The videos in the bottom part of Figure 3 are YouTube tutorials for the teacher, which contain rules of use to facilitate their use, despite being as intuitive and accessible as possible. It consists of six videos in total, which are in the corresponding picture itself [30].

Evaluación del Profesor	Evaluación propia Alumno	Evaluación Alumno/Compañero	Evaluación Alumno/Grupos
Porcentaje(%): <input type="text"/>	Porcentaje(%): <input type="text"/>	Porcentaje(%): <input type="text"/>	Porcentaje(%): <input type="text"/>
<b>Lista de alumnos:</b>	<b>Práctica 1:</b>	<b>Práctica 2:</b>	<b>Práctica N:</b>
Alumno 1: <input type="text" value="A1"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Alumno 2: <input type="text" value="A2"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Alumno 3: <input type="text" value="A3"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Alumno N: <input type="text" value="AN"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
			<b>Nota FINAL:</b>
			<input type="text"/>
			<input type="text"/>
			<input type="text"/>
			<input type="text"/>
<input type="button" value="Guardar"/>			

Figure 3. Interface for the teaching staff. Source: own compilation (the content of the figure appears in Spanish language because it is directly extracted from the platform). [30].

Consecutively, the program requires the allocation of evaluation percentages for the single agents and tasks to obtain a weighted qualification. From all of these and other student evaluations, it allows dumping an individual and group pdf document with all their marks. See the summary in Figure 3.

The interface of the student has simpler content and use (Figure 4); it is totally intuitive, so as to avoid issues in the event of not having a good command of IT, even if the student was a digital native.

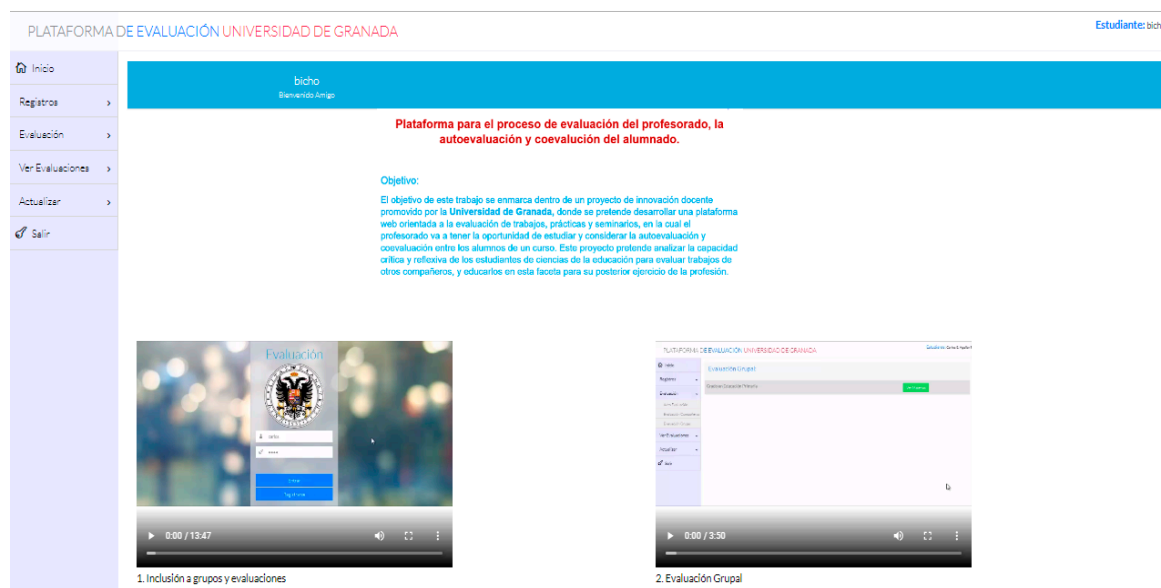


Figure 4. Start of PLEVALUA in the role of student. Source: taken from PLEVALUA [30].

Even so, despite its simplicity, it has three tutorial videos on its management, whose location is as follows and is accessed from the home interface [30].

See Figure 5 for the summary content of the student interface, which can be accessed after teacher authorization or registration.

Evaluación del Profesor	Evaluación propia Alumno	Evaluación Alumno/Compañeros	Evaluación Alumno/Grupos
Porcentaje(%): <input type="text"/>	Porcentaje(%): <input type="text"/>	Porcentaje(%): <input type="text"/>	Porcentaje(%): <input type="text"/>
Práctica 1: <input type="text"/>	Práctica 1: <input type="text"/>	A1 <input type="text"/> A2 <input type="text"/> A3 <input type="text"/>	G1 <input type="text"/> G2 <input type="text"/> G3 <input type="text"/>
Práctica 2: <input type="text"/>	Práctica 2: <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>
↓	↓	↓ ↓ ↓	↓ ↓ ↓
Práctica N: <input type="text"/>	Práctica N: <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>
Nota Bloq.: <input type="text"/>	Nota Bloq.: <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/> <input type="text"/>
<input type="button" value="Guardar"/>		Nota FINAL: <input type="text"/>	

Figure 5. Interface for students. Source: own compilation (the content of the figure appears in Spanish language because it is directly extracted from the platform). [30]

The platform is accessible from computers and other mobile devices, such as smartphones or tablets with an internet connection. It also allows the evaluation recording in paper format and its subsequent introduction, to avoid the digital divide, although all students have mobile devices and even laptops in the classroom, as well as individual internet connection or access via Wi-Fi.



### 3.2. Platform Validation

Regarding the use of PLEVALUA by students, there have not been any complications reflected. Through verbal comments and especially through the platform's suggestion box, favourable reviews concerning the combined evaluation model have been mostly compiled: "I think it is very appropriate, it was about time that students could also participate in their own and external progress because we know how everything develops in the classroom, more than the teachers themselves" (ES6) corresponding to the category of positive assessment. Positive reviews in the category "multiple evaluation assessment" (MEA) occurred to a greater extent (55.5% of the statements), and likewise in a smaller proportion (35.5%) for the category the "use of the platform assessment" (AUP):

- "I like this, (...) the platform allows you to evaluate at any time and anonymously so that other students do not know because it is something that requires privacy" (ST13);
- "(...) does not pose any difficulty once you look at the tutorials" (ST4);
- "(...) the tutorials are better than the written explanation for teachers and students that other platforms offer" (TE7).

The declared problems usually emanate from the "critique of multiple evaluation" (CME), both its own, for its extreme subjectivity, as well as others, for the generated inbreeding, and only testimony on the "critique of the use of platform" (CUP). They just represent 10% of the statements together, and it should also be emphasized that some respond in a positive manner, barely critical or negative:

- "(...) some classmates overestimated their effort and their mark, and also that of their friends, but it has been gradually controlled by the teacher's emphasis" (ST10);
- "(...) delivering evaluation to students can bring about issues, both about their evaluation and that of others, even if it is very modern (...)" (TE5);
- "I regard the employment of a platform for the serious task of evaluating as inappropriate (...)" (ST9).

In general, positive responses predominate from both the combination of evaluations (MEA) as well as the use of the platform (AUP) for this purpose (90%). There are no differences in this aspect between one group and the other (teaching staff: 88.45% and students: 92.5%). On the contrary, the criticism is a testimonial (10%), without a difference between those of one group and the other: (teaching staff: 11.55% and students: 8.5%). Given the anonymous nature of the dump of these data, it has not been possible to calculate data on the influence of other independent variables (such as sex, age, etc.); they have rather been taken as two cases for every single group.

Conversely, for the quantitative assessment of the platform through the scale, after its presentation and corresponding training, it offers optimal results as a whole around the perception of the suitability of the multi-evaluation method, whose items and partial scores are seen in Table 1.

The ratings are high for both (teachers) and others (students) in their commitment to the combined evaluation model (item 1), which is consistent with the new university education (item 11), so much that they are in favour of implementing it in their lessons (item 3).

Specifically, at the time of materializing the above, it seems appropriate to evaluate the members of the operational working groups who know the performance of each member in the same group (item 5), and even that of the rest of colleagues after presenting the produced work (item 7) in class.

Although, in these cases (item 5 and 7) not only does the average decrease concerning the first, which are the more general ones (items 1, 3 and 11), but also the fashion in the teaching group that goes beyond the maximum value of the item group before the value 3, although not so in the student group.

This modal value 3 also acquires the scores offered before the self-assessment (item 9), both for teachers and students (only case), and the average falls considerably ( $M = 2.50$  and  $2.90$ , respectively), observing greater dispersion of the data ( $sd = 0.96$  and  $0.95$ , respectively). Although the averages and the trends continue pointing towards a high value, it is presented as the typology that generates more uncertainties and controversy. Furthermore, it is the only case in which it is possible to verify

generalisable differences, that is to say significant differences ( $t = 3.68, p = 0.04$ ) between the teaching and student groups.

**Table 1.** Teachers' and students' opinion about multiple evaluation.

	M and Sd		Mo and %		t Student p-Value
	TE	ST	TE	ST	
1. Does it seem appropriate to you to combine traditional (hetero) assessment with peer (co) and self (self) assessment?	3.70 0.85	3.85 0.78	4 50	4 85.02	t = 2.54 p = 0.85
3. The teacher must begin to share the evaluation process with the students themselves	3.60 0.81	3.80 0.75	4 70.25	4 70.15	t = 1.35 p = 0.61
5. The operational co-workers of each practice group should also be evaluators ...	3.55 0.61	3.75 0.84	3 61.11	4 58.50	t = 3.21 p = 0.09
7. The practices presented in class by each group must be evaluated by the rest of the groups ...	3.55 0.69	3.70 0.77	3 55.55	4 63.88	t = 0.94 p = 0.08
9. Each student is a builder of their own learning and must also be an evaluator of it ...	2.50 0.96	2.90 0.95	3 40.80	3 50%	t = 3.68 p = 0.04
11. Do you think the current assessment needs to adapt to the current EHEA where the student is most active?	3.85 0.49	3.90 0.64	4 75.33	4 85.85	t = 1.82 p = 0.16
TOTAL average-mode dispersion	3.46 0.74	3.65 0.79	4 - 3		

Source: own compilation. Abbreviations: (European higher education area (EHEA)).

According to the calculations of the correlation coefficient (Pearson for students and Spearman for teachers), there are frequent relevant and intense correlations ( $r > 0.6, p < 0.05$ ) between items, but not with self-assessment ( $r = 0,35, p = 0.04$  for students and  $\rho = 0.25, p = 0.02$  for teachers).

The differences between the participants were not significant ( $p < 0.05$ ), neither regarding sex and contact with digital platforms, according to the calculation of Student's t, nor age and teaching experience, according to the ANOVA calculation (it should be noted that it was a sample of relatively young teachers not having extensive experience). A uniform or well-configured pattern can be used for multi-evaluation.

Once the unanimous and determined commitment to multi-evaluation is identified, it is necessary to analyse the value and usefulness assigned to the use of the platform as a facilitator of the registration and computation of all the evaluations of all agents involved in the didactic act for when it is performed.

The valuation is even more favourable than the previous one. All the means to all the items about the contemplation of each evaluation typology and its combination are high ( $M > 3.60$ ) and the model reaches the maximum estimated value (Table 2). Also, no significant differences ( $p > 0.05$ ) are found among the participating groups, as can also be seen in the table.

On the previous dimension, there is high consistency between the answers offered and, thus, some robustness on the pattern drawn on PLEVALUA, as evidenced by the very frequent relationship between answers. The correlations between the facilities of the platform proliferate to complete complementary evaluations, with a predominance of direct and intense relationships ( $r > 0.6$  for students and  $\rho > 0.6$  for teachers) and significant ( $p < 0.05$ ).

Furthermore, there are no significant differences ( $p > 0.05$ ) by sex, familiarity with virtual platforms, age, or university teaching experience, according to the calculations of Student's t and ANOVA.

As the scale has been designed to first identify the perception on specific issues of the diverse evaluation (odd items, first block of the scale, cf. Table 1) and then on their views on the functionality of such evaluations through the platform (items of Table 2), the odd items are linked to their consecutive pairs. In this way, it is possible to relate them to each other, with the hypothesis that odd scores correlate with those of peers.

**Table 2.** Teaching and student perceptions about the usefulness of PLEVALUA.

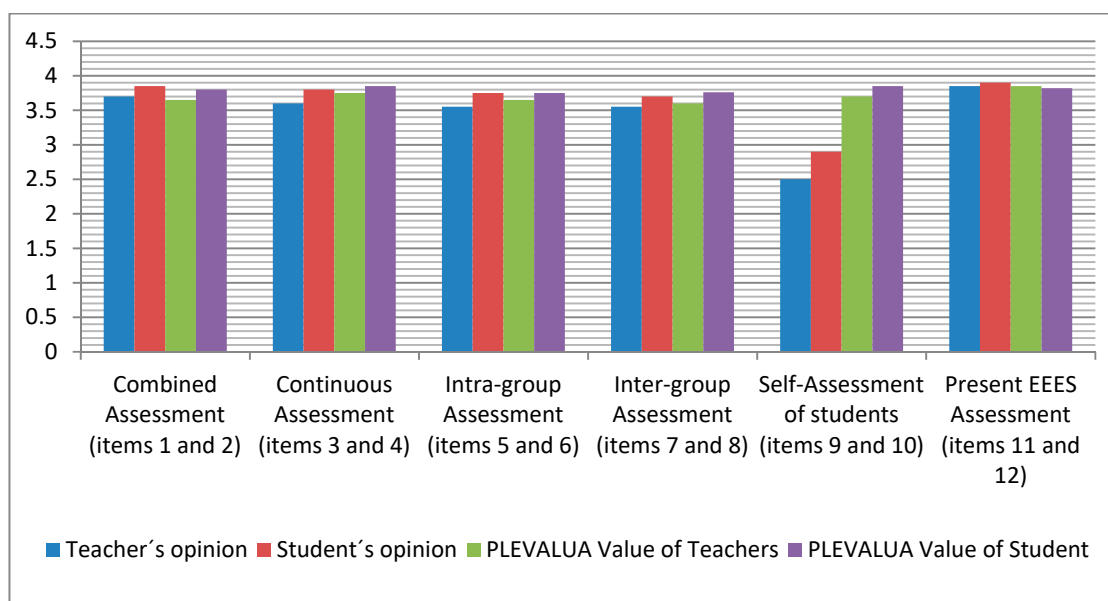
	M and Sd		Mo and %		t Student <i>p</i> -Value
	DO	ES	DO	ES	
2. Does the platform enable or favour this combined assessment in a university class group?	3.65 0.62	3.80 0.53	4 65.50	4 80.20	t = 1.23 <i>p</i> = 0.45
4. Do you think that the continuous evaluation by the teacher through the platform is viable?	3.75 0.61	3.75 0.42	4 61.75	4 84.54	t = 2.20 <i>p</i> = 0.86
6. Do you think the platform makes the evaluation of colleagues in the same work group viable?	3.65 0.61	3.76 0.46	4 55.40	4 78.56	t = 4.23 <i>p</i> = 0.77
8. Do you think it allows the continuous evaluation of other work groups within the classroom?	3.60 0.60	3.70 0.49	3 57.33	4 79.89	t = 0.94 <i>p</i> = 0.56
10. Do you think it allows the student's continuous self-assessment during their subject practices?	3.70 0.51	3.85 0.40	4 58.50	4 83.33	t = 1.89 <i>p</i> = 0.68
12. Do you think that the evaluation platform involves a current evaluation according to the EHEA in class?	3.85 0.51	3.82 0.45	4 54.45	4 81.05	t = 0.95 <i>p</i> = 0.06
TOTAL average-mode dispersion	3.70 0.56	3.81 0.48	4		

Source: own compilation.

This has been the case in all pairs of scores (1–2, 3–4, 5–6, 7–8 and 11–12), both for the sample of teachers and students, as indicated by the lack of significant differences between them, according to Student's *t* ( $p \geq 0.05$ ), and the manifestation of intense and direct correlations, regarding Pearson's *r* for students ( $r > 0.6$ ) and Spearman's rho for teachers ( $\rho \geq 0.6$ ) and lack of significant differences ( $p > 0.05$ ).

There is a curious exception with the pair of items 9–10 referred to in the self-assessment, where both groups, although especially that of the teaching staff, do not show enough conviction and confidence in the self-assessment (item 9:  $M_{TE} = 2.50$  and  $M_{ST} = 2.90$ ). However, they do admit that the platform allows this option adequately (item 10:  $M_{TE} = 3.70$  and  $M_{ST} = 3.85$ ), which statistically implies not only the absence of correlation between the responses of both items, according to Pearson's *r* calculations with the student data ( $r = 0.29$ ,  $p = 0.05$ ) and Spearman's rho ( $\rho = 0.13$ ,  $p = 0.04$ ), but there exist significant differences between them, according to the Student's *t* calculation ( $t = 3.61$ ;  $p = 0.00$ ).

The following graphic in Figure 6 visualizes the described scenario.

**Figure 6.** Average scores for opinions on evaluation method and PLEVALUA.

#### 4. Discussion and Conclusions

It is appropriate at the time of the praxis of university teaching to put the theory that supports different evaluation modalities into action in the classrooms. All authors who have studied the subject in recent times encourage and support this idea, which does not always translate into educational praxis [1].

The active role of students in their learning construction has to be endorsed by the feasibility and participation in their own evaluation, as an absolute guarantor of reorientation of efforts and learning [19,24]; as well as the evaluation of their peers, who are familiar to their daily work and continuous learning [3,12,19]; and, of course, with the continuous supervision of the teacher [29], as the utmost responsible for their own evaluation process and that of the others, so that the real accreditation of the progress of the students takes place [10,11,31].

Perhaps it is lost in objectivity of the scores (if the exclusive evaluation of the teacher is objective), but the truth is that at present learning is not conceived as something objective (but it depends on the construction of each individual). It is also needless to mention that it is lost in the ease of the process, but learning is at the end of the day a complex process.

Except for the difficulties, it is all about the gains. Not only in the learning process itself and the teaching process but also the democratization of the teaching and involvement of the students. Despite this, certain reluctance from the teaching group and curiously that of the students has been observed in the research, although with less intensity, towards the assumption of self-evaluation, unparalleled with the rest of the evaluation modalities contemplated, more demanded by the students [9] and teachers [27,31]. Other research has revealed the benefits of self-assessment, but they have not estimated its acceptance by educational agents.

What is evident is the need for such a combination but also its drawbacks [6,10], as there are no appropriate tools for multi-evaluation; unlike other types of teaching planning platforms, teaching resources, self-learning, etc., which are well received by university teachers [27].

Thus, PLEVALUA has been presented and validated in this work, whose location, description and validation have been made explicit. There is no discussion about other works since the field of research and contribution in this line is poorly supported and hence the shortage of this type of experience and products and its variation process under the assumed research design.

However, there is evidence of the use of rubrics and evaluation rubrics [26], mainly in Excel spreadsheet format or similar to help to weight and calculate scores based on partial scores. Such formulas cannot become solvent synchronous teaching tools such as interactive and data storage platforms.

Therefore, without prejudice to continuing to investigate in the previous topic of the advantages and disadvantages of the combined multiple evaluations (hetero-evaluation, self-evaluation, and co-evaluation), the development of technological products will be welcome to the corpus of knowledge of this matter, generally, to undertake the evaluation of all, facilitating their registration and final calculation.

Currently, apart from continuing to perfect PLEVALUA to make it more eloquent and friendly as well as more universally accessible (for handicapped individuals), it is being simplified in order to enable the launch of an app version.

Furthermore, it also needs to be validated by students in years other than second and fourth, as well as by students from other degree and postgraduate disciplines. As for the faculty, PLEVALUA needs to be tested not only by new teachers but also by older teachers with more experience, and less familiarity with didactic digital platforms, to verify whether they have the same capability of usage towards the evaluating platform.

However, what should be pursued is not so much extending the use of this particular platform, but the creation of other more contextualised tools within institutions (according to their possibilities and demands), which foster in-person and distance assessment. These platforms should be validated by their users: teachers and students, not just in higher education but also in secondary education.

**Author Contributions:** Conceptualization, A.R.F.; methodology, A.R.F. and L.A.B.; software, A.R.F. and L.A.B.; validation, A.R.F., J.L.G.O. and A.R.F.; formal analysis, A.R.F. and J.L.G.O.; investigation, A.R.F., A.N.R. and J.L.G.O.; resources, A.R.F. and J.L.G.O.; data curation, A.R.F. and A.N.R.; writing—original draft preparation, A.R.F., A.N.R. and J.L.G.O.; writing—review and editing, A.R.F., L.A.B., A.N.R.; visualization, A.R.F. and L.A.B.; supervision, A.R.F. and A.N.R.; project administration, A.R.F., L.A.B., A.N.R. and J.L.G.O.; funding acquisition, A.R.F., L.A.B., A.N.R. and J.L.G.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research is the result of a teaching innovation project called “PLEVALUA: Combined Assessment Platform: hetero-evaluations, self-assessments, and co-evaluations of university students” funded by the Quality, Innovation and Prospective Unit of the University of Granada, Spain (code 16-78; academic year 2017/18).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Assessment scale of the multiple evaluations and the PLEVALUA platform.

Value From 1 (Total Disagreement) to 4 (Total Agreement)	1	2	3	4
1. Does it seem appropriate to combine traditional (hetero), peer (co) and self (self) assessments				
2. Does the platform enable or favour this combined assessment in a university class group?				
3. The teacher must begin to share the evaluation process with the students themselves				
4. Do you think that the continuous evaluation by the teacher through the platform is viable?				
5. The operational co-workers of each practice group should also be evaluators?				
6. Do you think the platform makes the evaluation of colleagues in the same work group viable?				
7. The practices presented in class by each group must be evaluated by the rest of the groups?				
8. Do you think it allows the continuous evaluation of other work groups within the classroom?				
9. Each student is a builder of their own learning and must also be an evaluator of it?				
10. Do you think it allows the student’s continuous self-assessment during their subject practices?				
11. Do you think the evaluation methodology is consistent with the current EHEA where the student is more active?				
12. Do you think the evaluation platform involves a current evaluation according to the EHEA in class?				

## References

- Pérez, R. Quo vadis, evaluación? Reflexiones pedagógicas en torno a un tema tan manido como relevante. *Rev. Investig. Educ.* **2019**, *34*, 13–30. [CrossRef]
- Moro, A.I. La evaluación de las técnicas de evaluación en la enseñanza universitaria: La experiencia de macroeconomía. *Cult. Educ.* **2016**, *28*, 843–862. Available online: <https://dialnet.unirioja.es/servlet/articulo?codigo=5738670> (accessed on 11 April 2019).
- Barriopedro, M.; de Subijana, L.; Gómez Ruano, M.A.; Rivero, A. La coevaluación como estrategia para mejorar la dinámica del trabajo en grupo: Una experiencia en Ciencias del Deporte. *Rev. Complut. Educ.* **2016**, *27*, 571–584. [CrossRef]
- Gallego Noche, B.; Quesada, M.A.; Gómez Ruiz, J.; Cubero, J. La evaluación y retroalimentación electrónica entre iguales para la autorregulación y el aprendizaje estratégico en la universidad: La percepción del alumnado. *REDU* **2017**, *15*, 127–146. [CrossRef]
- Ibarra, M.S.; Rodríguez Gómez, G. Modalidades participativas de evaluación. Un análisis de la percepción del profesorado y de los estudiantes universitario. *Rev. Investig. Educ.* **2014**, *32*, 339–361. [CrossRef]
- Rodríguez Gómez, G.; Ibarra, M.S.; García Jiménez, E. Autoevaluación, evaluación entre iguales y coevaluación: Conceptualización y práctica en las universidades española. *REINED* **2013**, *11*, 198–210. Available online: [File:///C:/Users/T101/Downloads/Dialnet-AutoevaluacionEvaluacionEntreIgualesYCoevaluacion-4734976%20\(2\).pdf](File:///C:/Users/T101/Downloads/Dialnet-AutoevaluacionEvaluacionEntreIgualesYCoevaluacion-4734976%20(2).pdf) (accessed on 1 February 2019).
- Ibarra-Sáiz, M.S.; Rodríguez Gómez, G. EvalCOMIX®: A web-based programme to support collaboration in assessment. In *Smart Technology Applications in Business Environments*; Issa, T., Kommers, P., Issa, T., Isaías, P., Issa, T.B., Eds.; IGI Global: Hershey, PA, USA, 2017; pp. 249–275. [CrossRef]



8. Lukas, J.F.; Santiago, K.; Lizasoain, L.; Etxebarria, J. Percepciones del alumnado universitario sobre su evaluación. *Bordón* **2017**, *69*, 103–122. Available online: <https://recyt.fecyt.es/index.php/BORDON/article/view/43843> (accessed on 5 February 2020).
9. Rodríguez Espinosa, H.; Retrespo, L.F.; Luna, G. Percepción del estudiantado sobre la evaluación del aprendizaje en la educación superior. *Educare* **2016**, *20*, 1–17. [[CrossRef](#)]
10. Gallego Ortega, J.L.; Rodríguez Fuentes, A. Alternancia de roles en la evaluación universitaria: Docentes y discentes evaluadores y evaluado. *REDU* **2017**, *15*, 349–366. [[CrossRef](#)]
11. Arnáiz, C.M.; Bernardino, A.C. Los resultados de los estudiantes en un proceso de evaluación con metodologías distintas. *Rev. Investig. Educ.* **2013**, *31*, 275–293. [[CrossRef](#)]
12. Pascual-Gómez, I.; Lorenzo-Llamas, E.M.; Monge-López, C. Análisis de validez en la evaluación entre iguales: Un estudio en educación superior. *RELIEVE* **2015**, *21*, 1–17. [[CrossRef](#)]
13. Quesada, V.; García-Jiménez, E.; Gómez-Ruiz, M.A. Student Participation in Assessment Processes. In *Learning and Performance Assessment*; Khosrow, M., Ed.; IGI Global: Hershey, PA, USA, 2020; pp. 1226–1247. Available online: <https://www.igi-global.com/chapter/student-participation-in-assessment-processes/237579> (accessed on 1 May 2020).
14. Mok, M.M.; Lung, L.; Cheng, D.P.W.; Cheung, H.P.; Ng, M.L. Self-assessment in Higher Education: Experience in using a metacognitive approach in five case studies. *Assess. Eval. High. Educ.* **2017**, *31*, 415–433. [[CrossRef](#)]
15. Yan, Z.; Brown, G.T.L. A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assess. Eval. High. Educ.* **2017**, *42*, 1247–1262. [[CrossRef](#)]
16. Yan, Z. Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assess. Eval. High. Educ.* **2019**, *44*, 224–238. [[CrossRef](#)]
17. Boud, D.; Falchikov, N. *Rethinking Assessment in Higher Education*. *Learning Longer Term*; Routledge: London, UK, 2017.
18. Boud, D.; Molloy, E. *Feedback in Higher and Professional Education*; Routledge: London, UK, 2013.
19. Carrizosa, E.; Gallardo, J.I. *Autoevaluación, Coevaluación y Evaluación de los aprendizajes* In *III Jornadas sobre docencia del Derecho y TIC*; Barcelona, E., Khosrow-Pour, M., Eds.; Information Resources Management Association: New York, NY, USA, 2020; Available online: [http://www.uoc.edu/symposia/dret\\_tic2012/pdf/4.6.carrizosa-esther-y-gallardo-jospdf](http://www.uoc.edu/symposia/dret_tic2012/pdf/4.6.carrizosa-esther-y-gallardo-jospdf) (accessed on 10 October 2019).
20. López Lozano, L.; Solís Ramírez, E. Con qué evalúan los estudiantes de magisterio en formación. *Campo. Abierto*. **2016**, *35*, 55–67. Available online: <file:///C:/Users/T101/Downloads/Dialnet-ConQueEvaluanLosEstudiantesDeMagisterioEnFormacion-5787079.pdf> (accessed on 15 May 2019).
21. Ion, G.; Cano, E. El proceso de implementación de la evaluación por competencias por competencias en la Educación Superior. *REINED* **2011**, *9*, 246–258. Available online: <http://reined.webuviges/index.php/reined/article/view/128> (accessed on 25 June 2019).
22. Tejada Fernández, J.; Ruiz Bueno, C. Evaluación de competencias profesionales en educación superior: Retos e implicaciones. *Educación XXI* **2015**, *19*, 17–38. [[CrossRef](#)]
23. Morrell, L.J. Iterated assessment and feedback improves student outcomes. *Stud. High. Educ.* **2019**, *44*, 105–123. [[CrossRef](#)]
24. Rodríguez Gómez, G.; Ibarra, M.S.; Gómez Ruiz, M.Á. e-Autoevaluación en la universidad: Un reto para profesores y estudiante. *Rev. Educ.* **2011**, *356*, 401–430. Available online: <http://www.educacionyfp.gob.es/revista-de-educacion/numeros-revista-educacion/numeros-anteriores/2011/re356/re356-17.html> (accessed on 2 July 2019).
25. García Jiménez, E. La evaluación del aprendizaje: De la retroalimentación a la autorregulación. El papel de las tecnologías. *RELIEVE* **2015**, *21*, 1–24. [[CrossRef](#)]
26. Saiz, M.; Bol, A. Aprendizaje basado en la evaluación mediante rúbricas en educación superior. *Suma Psicológica* **2014**, *21*, 28–35. Available online: <http://www.sciencedirect.com/science/article/pii/S0121438114700049> (accessed on 1 July 2019). [[CrossRef](#)]
27. De Pablos, J.; Colás, M.; López Gracia, A.; García-Lázaro, I. Los usos de las plataformas digitales en la enseñanza universitaria. Perspectivas desde la investigación educativa. *Rev. Docencia Univ.* **2019**, *17*, 59–72. [[CrossRef](#)]
28. De Benito, B.; Salinas, J.M. La investigación basada en diseño en Tecnología Educativa. *Rev. Interuniv. Investig. Tecnol. Educ.* **2016**, *44*–59. [[CrossRef](#)]

29. Pressman, S. *Ingeniería del Software: Un Enfoque Práctico*, 5th ed.; McGraw-Hill/Interamericana de España: Madrid, España, 2002.
30. Plataforma de Evaluación UGr. Available online: [http://www.linyadoo.com/plevalua\\_ug](http://www.linyadoo.com/plevalua_ug) (accessed on 14 May 2019).
31. Quesada, V.; Rodríguez Gómez, G.; Ibarra, M.S. Planificación e innovación de la evaluación en educación superior: La perspectiva del profesorado. *Rev. Investig. Educ.* **2017**, *35*, 53–70. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).