





## University teachers' perceptions regarding the quality of tasks assessing learning outcomes

*Percepción del profesorado universitario sobre la calidad de las tareas de evaluación de los resultados de aprendizaje*

*Percepção dos docentes universitários sobre a qualidade das tarefas de avaliação dos resultados da aprendizagem*

*大学教师对学习结果评估质量的想法*

*تصور أعضاء هيئة التدريس بالجامعة لجودة مهام تقييم مخرجات التعلم*

Ibarra-Sáiz, M<sup>a</sup>. Soledad<sup>(1)</sup> , Lukas-Mujika, José-Francisco<sup>(2)</sup> , Ponce-González, Nicolás<sup>(1)</sup> ,  
Rodríguez-Gómez, Gregorio<sup>(1)</sup> 

<sup>(1)</sup> UNESCO Chair on Evaluation and Assessment, Innovation and Excellence in Education. University of Cadiz, Spain;

<sup>(2)</sup> University of the Basque Country /Euskal Herriko Unibertsitatea, Spain

### Abstract

Assessment of whether learning outcomes have been achieved requires teaching, assessment and learning to be constructively aligned, highlighting the importance of designing assessment tasks that meet the necessary quality conditions to strengthen student learning. This study was carried out to analyse university lecturers' perception of their design characteristics in the assessment tasks as part of their evaluative practice. The study followed a mixed methodology (exploratory sequential design) using the RAPEVA questionnaire -Self-report from teaching staff on their practice in learning outcome assessment. This questionnaire collected opinions from 416 teachers working at six public universities in various Spanish autonomous regions. The transparency, through information provided to the students, and the depth of the tasks are two aspects often mentioned by the teachers. On the other hand, feedback or participation from students in assessment processes are aspects which teachers consider less important. This detects differences in perception depending on the university, the field of knowledge and how secure and satisfied the teachers feel regarding the assessment system. In accordance with the results, future lines of research are suggested that favour greater understanding of evaluative practices in higher education.

**Keywords:** Assessment, Learning, Higher Education.

### Resumen

Evaluar la consecución de los resultados de aprendizaje exige un alineamiento constructivo entre enseñanza, evaluación y aprendizaje, en el que se destaca la importancia del diseño de tareas de evaluación que cumplan con las condiciones de calidad suficientes para que sean potenciadoras del aprendizaje del estudiante. Este estudio se ha realizado con la finalidad de analizar la percepción que tiene el profesorado universitario de las características de las tareas de evaluación que diseña en su práctica evaluativa. El estudio ha seguido una metodología mixta (diseño secuencial exploratorio) utilizando el cuestionario RAPEVA-Autoinforme del profesorado sobre su práctica en la evaluación de resultados de aprendizaje. Se ha recabado la opinión de 416 profesores de seis universidades públicas de diferentes comunidades autónomas españolas. La transparencia, a través de la información que se facilita a los estudiantes y la profundidad de las tareas son los dos aspectos más destacados por el profesorado. En cambio, la retroalimentación o la participación del estudiantado en los procesos de evaluación son aspectos menos considerados por parte del profesorado. De destacar las diferencias de percepción detectadas en función de la universidad, el ámbito de conocimiento y el grado de seguridad y satisfacción con el sistema de evaluación. En consonancia con los resultados, se ofrecen futuras líneas de investigación que favorezcan una mayor comprensión de las prácticas evaluativas en educación superior.

**Palabras clave:** Evaluación, Aprendizaje, Educación Superior.

Received/Recibido

Feb 13, 2023

Approved /Aprobado

Apr 12, 2023

Published/Publicado

Jun 29, 2023

## Resumo

Avaliar a consecução dos resultados de aprendizagem exige um alinhamento construtivo entre ensino, avaliação e aprendizagem, no qual se destaca a importância de conceber tarefas de avaliação que cumpram as condições de qualidade suficientes para potenciarem a aprendizagem do estudante. Este estudo foi realizado com a finalidade de analisar a percepção dos professores universitários sobre as características das tarefas de avaliação que concebem na sua prática de avaliação. O estudo seguiu uma metodologia mista (design sequencial exploratório) utilizando o questionário RAPEVA-Autorrelato dos professores sobre a sua prática na avaliação de resultados da aprendizagem. Recolheu-se a opinião de 416 professores de seis universidades públicas de diferentes comunidades autónomas espanholas. A transparência, através da informação fornecida aos alunos, e a profundidade das tarefas são os dois aspetos mais salientados pelo pessoal docente. Por outro lado, o feedback ou a participação dos estudantes nos processos de avaliação são aspetos menos considerados pelos professores. É de salientar as diferenças de percepção detetadas em função da universidade, do âmbito de conhecimento e do grau de segurança e satisfação com o sistema de avaliação. De acordo com os resultados, oferecem-se linhas de investigação futuras que favoreçam uma maior compreensão das práticas de avaliação no ensino superior.

**Palavras-chave:** Avaliação, Aprendizagem, Ensino Superior.

## 摘要

对学习结果的评估需要在教学、评估及学习间实现建设性对齐，因此设计符合质量标准且能推动学生学习的评估任务就变得尤为重要。该研究的主要目的是分析大学教师对评估实践中的评估任务特点有哪些看法。研究遵循混合方法（探索性顺序法），使用 RAPEVA 问卷，教师对学习成果评估实践的自我反馈报告，得到来自西班牙各自治区 6 所公立大学共 416 名教师的意见。其中给学生提供信息的透明度和任务深度是教师观点中最突出的两个方面。相反，学生的反馈和参与是教师评估过程中考虑最少的两项内容。值得强调的还有来自不同大学的教师的态度差异、知识领域、对评估系统的安全感和满意值这几个因素。该研究的成果为未来的研究提供了方向，让我们对高等教育评估实践的理解更进一步。

**关键词:** 评估、学习、高等教育

## ملخص

يتطلب تقييم تحقيق مخرجات التعلم المواءمة البناء بين التدريس والتقييم والتعلم، مما يبرز أهمية تصميم مهام التقييم التي تلبى شروط الجودة الكافية لتعزيز التعلم الطلاب. أجريت هذه الدراسة بهدف تحليل التصور السائد بأن أعضاء هيئة التدريس بالجامعة لديهم خصائص مهام التقييم التي يصممونها في ممارستهم التقييمية. اتبعت الدراسة منهجية مختلطة (تصميم تسلسلي استكشافي) باستخدام حول ممارساتهم في تقييم نتائج التعلم. تم الحصول على رأي 416 أستاذًا من ست RAPEVA-Teacher استبيان التقرير الذاتي جامعات عامة في مختلف مجتمعات الحكم الذاتي الإسبانية. الشفافية، من خلال المعلومات المقدمة للطلاب، وعمق المهام هما الجانبان الأكثر إبرازًا من قبل المعلمين من ناحية أخرى، فإن التغذية الراجعة أو مشاركة الهيئة الطلابية في عمليات التقييم هي جوانب لا يأخذها أعضاء هيئة التدريس بعين الاعتبار. لإبراز الفروق في الإدراك المكتشفة حسب الجامعة ومجال المعرفة ودرجة الأمان والرضا عن نظام التقييم. تمشيا مع النتائج، يتم تقديم خطوط بحثية مستقبلية لتعزيز فهم أفضل لممارسات التقييم في التعليم العالي.

الكلمات الدالة: تقييم، التعلم، التعليم العالي

## Introduction

When designing assessment processes, teachers must make decisions on many curricular elements (Bearman et al., 2014, 2016). This paper focuses on the quality of assessment tasks as this is essential within the evaluative process.

This contribution explores and analyses the evaluative practice of university teachers,

working from their own perceptions and experience, specifically focusing on the type of assessment tasks they design, in other words, characteristics that help differentiate assessment tasks. It should be highlighted that this research falls within the broader context of the FLOASS Project (Ibarra-Sáiz & Rodríguez-Gómez, 2019). This project provides an action framework, supported by technologies that improve the assessment

(Technology Enhance Assessment (TEA) and Learning Analytics (LA)), that steer the design, implementation, monitoring and assessment of the Learning Outcomes (LO) which require students to demonstrate greater skills.

From different evaluative foci, papers by Carless (2015), Rodríguez-Gómez & Ibarra-Sáiz (2015) or Sambell et al. (2013) highlight the relevance of design in assessment tasks. Nevertheless, how do university teachers characterise the assessment tasks which they design in their evaluative practice? Is there a difference in the design of these assessment tasks that can be associated with characteristics such as their university, their field of knowledge, their gender, their years of experience or their perceived security and satisfaction with the assessment?

To answer the questions raised above, based on teachers' perceptions, this research aims to:

- analyse the characteristics of assessment tasks designed by university teachers;
- check whether differences in the assessment task design can be linked to certain contextual characteristics such as the university, field of knowledge, experience or satisfaction with the assessment system.

### The assessment tasks and their characteristics

Two questions need to be asked of assessment: What is the purpose of our assessment? What should we assess? Working from the answers to these two questions, we find various evaluative approaches in higher education. Based on contributions from authors such as Boud (2022), Ibarra-Sáiz et al. (2021) and Sambell et al. (2013), Table 1 provides a brief summary of the three fundamental approaches.

Table 1. Purpose of the assessment and the learning

<b>Evaluative purpose</b>	<b>What is the purpose of our assessment?</b>	<b>What should we assess?</b>
Assessment <i>of the</i> learning	To certify or inform the student and anyone else interested in the student's competence regarding the learning outcomes	How well students can apply the key concepts, knowledge, skills and attitudes related to the learning outcomes
Assessment <i>for the</i> learning	To allow teachers to determine the next steps to make progress in the student's learning	Each student's progress and learning needs in relation to the chosen learning outcomes
Assessment <i>as</i> learning	To guide and provide opportunities for each student to monitor and think critically about their learning and identify the next steps to take	What each student thinks about their learning, their strategies to support or challenge this learning and the mechanisms they use to adapt and make progress

Assessment tasks play a central role in any of these approaches. Firstly, the type of assessment task is related to the students' approach to their learning (Biggs & Tang 2011). Secondly, papers by Panadero et al. (2022) and López Gil et al. (2022) show that assessment activities strongly affect students' emotions and motivation. Finally, assessment tasks provide the evidence which teaching staff

can use to evaluate how well they have achieved learning outcomes, any possible modifications required to improve the teaching-learning process or guide the students on strategies to boost their progress and increase their learning.

For Ibarra-Sáiz, Rodríguez-Gómez & Boud (2021), a task is considered good quality if it is rigorous, credible, interesting and it promotes valuable learning for the student. This study focuses on four essential elements of tasks: transparency, depth, feedback and participation.

Transparency refers to information that must be given to the student, which Gore et al. (2009) conceptualised as support for the student. Consequently, the assessment task must be public, known and understood by the student. The student must recognise the quality of the products or actions that they must produce to achieve the expected learning outcomes; know what they are expected to learn and perform in the task; identify the assessment methods and instruments that are going to be used to assess their performance; and know what their role is going to be as an assessor through methods such as self-assessment, peer assessment and/or co-assessment. As mentioned by Yan & Boud (2022), to develop student learning, the assessment tasks must allow judgements on learning results not only from the teaching staff but also from the students themselves. To do this, students must be offered information and basic guidance that help them tackle the tasks successfully.

A good quality assessment task requires students to determine complex and coherent relationships between fundamental and significant concepts, which implies in-depth knowledge (Gore et al., 2009). In addition, one important change over the last decade involves an increase in authentic assessment (Boud 2020) by using complex tasks (Sambell et al., 2013), which implies using contextualised tasks and processes in professional practice. This type of task not only raises student motivation and commitment but also offers students the chance to solve real-life problems with greater meaning for their learning (Yan & Boud, 2022). In short, this refers to tasks that promote in-depth learning through investigation methods and critical and reflexive thinking (Ibarra-Sáiz & Rodríguez-Gómez, 2020).

Contributions by Henderson et al. (2018), Carless (2020), Lipnevich & Panadero (2021) and Boud & Dawson (2021) demonstrate the importance and relevance of feedback in the assessment process. Consequently, assessment task design must foresee how this feedback will be used and promoted, the role that students play in giving and receiving feedback and how students use the information provided.

Student participation in the assessment process can arise at different times and in different situations: from design to scoring, from writing the assessment criteria to designing and specifying the assessment methods and instruments. Papers by Falchikov (2005), Falchikov & Goldfinch (2000) and the latest contributions from Panadero & Alqassab (2019), Hortigüela Alcalá et al. (2019) and Quesada-Serra et al. (2019) state the positive effect of students' participation in their own learning process, by using evaluative methods such as self-assessment, peer assessment or co-assessment.

## Method

### *Research design*

A mixed methodology was chosen (qual->QUAN) using exploratory sequential design (Creswell, 2015). The first phase designed and validated the content of the RAPEVA questionnaire -Self-report from teaching staff on their practice in learning outcome assessment. In the second phase, in the final period of the 2020/21 academic year, perceptions were collected from social science teachers working at six universities.

### *The RAPEVA self-report questionnaire*

The RAPEVA self-report construction began with a literature review and subsequently, the group consensus method was used (Johnson & Morgan 2016) to validate the content, involving 22 judges in three iterations. The definition and specification of the different indicators was reviewed at the end of each iteration.

The RAPEVA self-report is based on linear combinations of the variables observed (Henseler 2021). In the evaluation made by each respondent on each of the self-report items, an essential role is played by the cognitive and behavioural aspects, and so these items built a formative index (Hair et al. 2022). Consequently, a generalised analysis of structural components (Hwang & Takane, 2015) was performed, obtaining adjustment measures (GFI=.89 and SRMR=.08) that are considered acceptable.

Regarding its structure, in the first part of the self-report, information is requested on

contextual aspects such as their university, their field of knowledge, years of experience or gender. The second presents 49 items in Likert scale format (0-5) structured into eleven dimensions (Table 2). It takes around 20 minutes to fill in the self-report.

As mentioned above, this study focuses on four dimensions related to the quality of the assessment tasks (TRA, PRO, RET and PAR). The results related to the assessment methods and instruments can be consulted in the paper by Ibarra-Sáiz et al. (2023).

Table 2. Structure of the RAPEVA self-report

Dimensions		# Items	Items
TRA *	Transparency	5	I01, I03 to I05, I35
CAE	Competences to be evaluated	6	I06 to I11
MOB	Observation instruments	6	I12 to I17
MEN	Survey instruments	4	I18 to I21
MDA	Documents and artifacts	9	I22 to I30
INE	Assessment instruments	4	I31 to I34
PRO*	Depth of the tasks	4	I36 to I38
RET*	Feedback	3	I39 to I41
PAR*	Participation	4	I42 to I45
FOR	Training being evaluated	2	I46, I47
SSE	Satisfaction with the assessment	2	I48, I49

### Participants

The questionnaire was sent in an online format to all teachers working on the master's courses for Communication, Education and Economics and Business at the six universities participating in the FLOASS project (<http://floass.uca.es>): University of Cádiz (UCA), Rovira i Virgili University (URV), University of Oviedo (UNIOVI), University of Valencia (UV), University of A Coruña (UDC) and University of the Basque Country (UPV/EHU). All the teachers received an

email from the coordinators, inviting them to use the link provided to fill in the self-report. Out of a sample of 2,400 teachers invited, a total of 626 teachers began to fill it in and 416 complete self-reports were received from teachers who taught on 63 different master's degrees, achieving a confidence level of 95%.

Table 3 presents the distribution of the 416 teachers depending on their university of origin, gender, field of knowledge, years of experience and degree of security and satisfaction with the assessment.

Table 3. Demographic characteristics

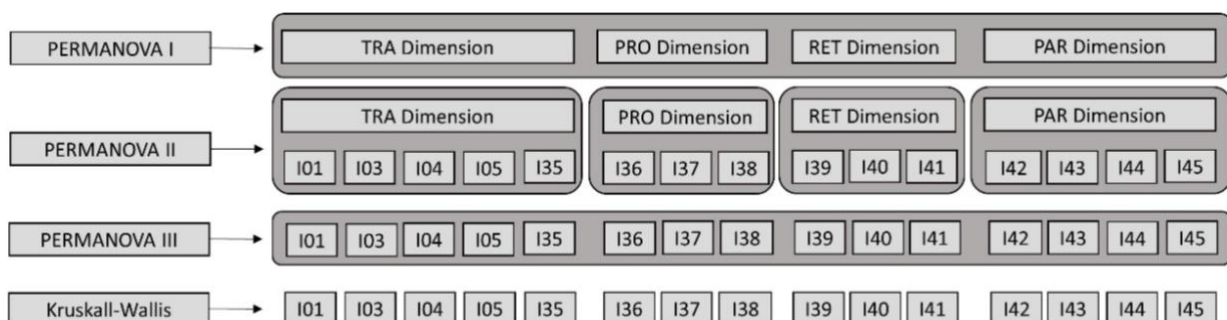
	Female		Male		Others		Total	
	n	%	n	%	n	%	n	%
<i>University</i>								
UCA	30	49.2	31	50.8	0	0	61	14.7
URV	18	60	12	40	0	0	30	7.2
UNIOVI	35	63.6	20	36.4	0	0	55	13.2
UV	36	51.4	33	47.1	1	1.4	70	16.8
UDC	45	54.2	37	44.6	1	1.2	83	20
UPV/EHU	61	52.1	56	47.9	0	0	117	28.1
<i>Field</i>								
COM	10	55.6	8	44.4	0	0	18	4.3
ECO	83	49.7	82	49.1	2	1.2	167	40.1
EDU	132	57.1	99	42.9	0	0	231	55.5
<i>Experience</i>								
<10 years	59	64.8	32	35.2	0	0	91	21.9
11-20 years	78	64.5	41	33.9	2	1.7	121	29.1
>20 years	88	43.1	116	56.9	0	0	204	49
<i>Satisfaction</i>								
Level 1	75	68.2	34	30.9	1	0.9	110	26.4
Level 2	104	59.1	72	40.9	0	0	176	42.3
Level 3	8	26.7	22	73.3	0	0	30	7.2
Level 4	38	38	61	61	1	1	100	24
<b>Total</b>	<b>225</b>	<b>54.1</b>	<b>189</b>	<b>45.4</b>	<b>2</b>	<b>0.5</b>	<b>416</b>	<b>100</b>

### Data analysis

An exploratory statistical analysis was performed to achieve the primary objective, and the differences between groups were analysed as the second analysis. As these are ordinal measurements that do not adjust to normality (K-S test,  $p < .001$ ), different non-parametric techniques have been used. The PERMANOVA-Permutational Multivariate Analysis of Variance (Anderson, 2017) was used at three different times (Figure 1). This technique has a similar aim to MANOVA, so it is also known as non-parametric MANOVA, although the difference lies in that it is based on the analysis of permutations on the distance

matrices to make the multivariate comparison. Firstly, the possible differences were contrasted between groups, comparing the four overall dimensions (PERMANOVA I). Secondly, to improve comprehension of the differences encountered, the elements making up each dimension were compared individually (PERMANOVA II). Finally, to go into greater depth on relationships between variables, a multivariate analysis was performed with all the items simultaneously (PERMANOVA III). Finally, each item was analysed individually using the Kruskal-Wallis H test. The analyses were performed using JASP (JASP Team, 2022) and R (R Core Team, 2021).

Figure 1. Comparative analysis process



## Results

Initially, the overall results were presented as obtained in the teachers' self-report answers, organised into the four characteristics of the assessment tasks being analysed: transparency, feedback, depth and participation. Subsequently, the results will be presented on differences found in the perceptions, using the university, field of knowledge, gender, experience and security and satisfaction with assessment as the comparison variables.

### *Characterisation of the assessment tasks*

Table 4 shows the central trend measurements (median and mean) and dispersion (typical deviation) for each of the self-report items grouped into the four aforementioned dimensions, which allows us to analyse the characteristics of the assessment tasks designed by the university teachers. For

the transparency dimension, there is a greater degree of agreement that the learning outcomes, criteria, assessment procedures and scoring systems for the subjects are public, as they appear in the teaching guides/programmes (I01,  $M=4.71$ ).

Regarding the depth aspect, the greatest degree of agreement appears when affirming that the assessment tasks imply the student's use of subject-relevant knowledge and content (I36,  $M=4.69$ ). Regarding the feedback dimension, the greatest degree of agreement is related to providing feedback to the students on their progress during the teaching-learning process (I39,  $M=3.98$ ). The means are very low in the participation dimension. The greatest degree of agreement, although low, is found when helping students to work together to specify elements of the assessment system (I42,  $M=2.67$ ).

Table 4. Central trend measurements and dispersion in the dimensions and items in the RAPEVA self-report

	Mdn	M	SD
<b>Transparency (TRA)</b>	4.2	4.18	.58
I01 The learning outcomes, criteria, assessment procedures and scoring system for the subject are public (they appear in the teaching guide/programme).	5	4.71	.73
I03 The assessment system provides information on what students must hand in or perform.	5	4.70	.66
I04 The assessment system provides information on the criteria and assessment instrument.	5	4.53	.81
I05 The assessment system provides information on the four assessment methods (self-assessment, peer assessment, co-assessment/shared assessment and assessment by teachers).	2	2.25	1.83
I35 I inform students and describe the assessment tasks that they will have to perform (using guides, appendices, transparencies, etc.).	5	4.70	.71
<b>Depth (PRO)</b>	4.67	4.44	.71
I36 The assessment tasks imply the student's use of subject-relevant knowledge and content.	5	4.69	.68
I37 The assessment tasks are set as a challenge for the student.	5	4.17	1.12
I38 The assessment tasks make it easier to apply knowledge and skills to situations or cases that are similar to what the student will find in the workplace.	5	4.45	.91
<b>Feedback (RET)</b>	3.33	3.30	1.31
I39 I provide feedback to the students regarding their progress during the teaching-learning process (by producing drafts offering an individual review or sessions/workshops on them, comparison with best practice or products, etc.).	4	3.98	1.26
I40 I make it easier for students to take part and provide their own feedback contrasting their progress with assessment instruments, comparison with best practice or products, etc. that I have given them previously.	4	3.32	1.63
I41 The students receive feedback on their progress from their classmates by means of peer review on drafts, oral presentations, etc.	3	2.60	1.91
<b>Participation (PAR)</b>	1.75	2.06	1.75
I42 I make it easier for students to work together on specifying some elements of the assessment system (products or learning actions to carry out, moments or delivery dates, proposal for assessment criteria, for assessment instruments, scoring system).	3	2.67	1.81
I43 The students self-assess their products or actions, either individually or in groups.	2	2.02	1.95
I44 The students assess their classmates' products or actions, either individually or in groups.	1	1.98	1.98
I45 Assessments and scores are given in a dialogue (between teachers and students) and as a consensus.	1	1.55	1.74

Figure 2 presents the position, dispersion and asymmetry of the scores from the teachers in the four dimensions. This overall perspective indicates that the greatest degree of agreement or frequency of use occurs in the dimensions of depth (M=4.44) and transparency (M=4.18)

and to a lesser degree in the dimensions of feedback (M=3.30) and participation (M=2.06). The independent results of each of these four dimensions are given below (Figure 3).

Figure 2. Boxplot corresponding to the four dimensions of the RAPEVA self-report

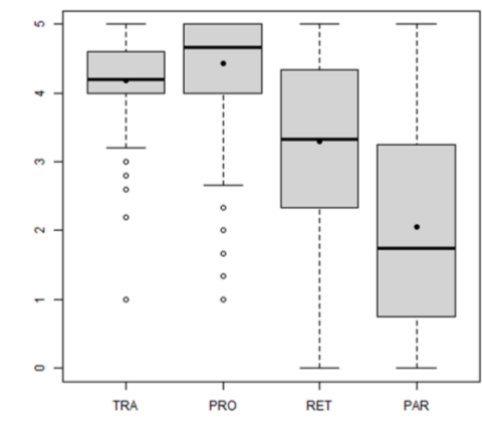
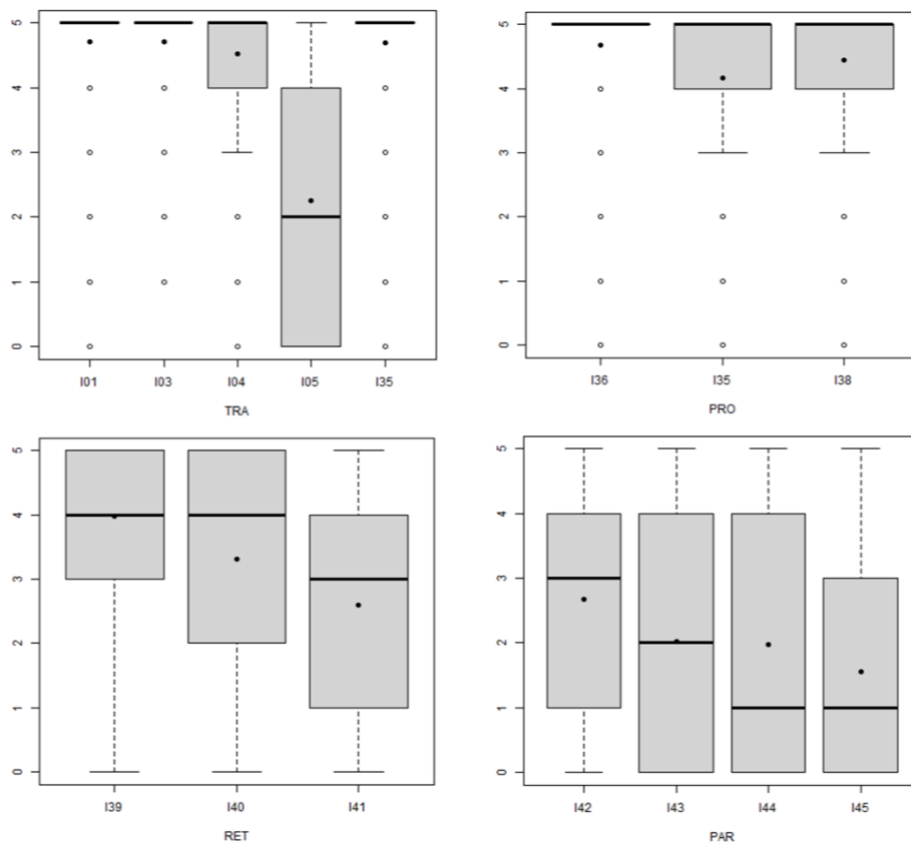


Figure 3. Boxplot of the items in the four dimensions





### ***Transparency in assessment tasks***

Figure 3 shows high, concentrated scores in the transparency (TRA) dimension items. In fact, Table 4 displays means over 4.5, except for item I05 (M=2.25). 95.1% of the teachers agree strongly with “the learning outcomes, criteria, assessment procedures and the scoring system for the subject are public and appear in the teaching guide or programme” (I01); in addition, 94.9% consider that “the assessment system provides information on the criteria and assessment instruments” (I03). In the same respect, 94.7% of teachers state that they believe they provide information to students and describe the assessment tasks that they have to perform and offer guides, appendices or slides to them, among other possible resources (I35). There is less agreement for item I04, as 91.9% consider that “The assessment system provides information on the criteria and assessment instrument.” Finally, it should be mentioned that there is a low agreement on whether the assessment system provides information on the various participative assessment methods (I05), such as self-assessment, peer assessment or co-assessment, where only 36% of teachers agree strongly with this statement.

### ***Depth of the assessment tasks***

In terms of the depth of the tasks (PRO), we can highlight the strong agreement from the teachers (95.4%) that the assessment tasks proposed to the students “imply the use of relevant knowledge and content to the subject” (I36). Next, 89.4% agree that these proposed tasks make it easier for students to apply knowledge and skills to similar situations or cases in the workplace (I38). Finally, 79.8% state that they agree the assessment tasks are set as a challenge for the students (I37).

### ***Feedback in the assessment tasks***

Regarding feedback, we see (Figure 3, RET) that there is greater dispersion and variability in these opinions compared to the previous two dimensions. In this case, 72.1% of teachers demonstrate that they give students feedback through strategies such as producing drafts where they provide an individual review,

carrying out placements or workshops or comparison with best practice or products (I39).

53.6% demonstrate that they make it easier for students to take part and give self-feedback, giving them assessment instruments beforehand, that they can use to contrast their progress against the standards explained in them, and also best practice or products to compare with their own (I40).

Finally, only 40.6% consider that students receive feedback from their own classmates through peer review of drafts, presentations, etc. (I41).

### ***Participation from the students in the assessment***

This participation dimension presents lower mean scores in all items (Table 4), all of them under 3. As seen in Figure 3 (PAR), there is greater variability in the scores, firstly highlighting that only 40.2% state they usually favour student collaboration when designing the assessment system elements such as specifying products or actions for evaluation, the times or dates that they are handed in, the actual assessment criteria, the design of instruments or the scoring system (I42). 30.8% state that they never manage to request this collaboration from the students.

Regarding implementation of the participative assessment methods, 48.8% state that they do not use self-assessment (I43) and over half the teachers (51.2%) say that they do not use peer evaluation (I44). In the case of co-assessment (I45), the percentage of teachers who do not use this strategy is 58.7%.

### ***Perceptive differences regarding the designed assessment tasks***

One of the research questions for this work refers to possible differences according to teachers' characteristics that might be useful for better comprehension of evaluative practice in higher education. To achieve this, comparative analysis, following a top-down strategy, started by analysing differences using four overall dimensions and then individually comparing each item from the self-report.

These analysis results are given in Tables 5 and 6.

An initial multivariate analysis (PERMANOVA I) considered the four dimensions to be dependent variables. These results are presented in Table 5. Significant differences were found depending on the university ( $F_{5, 415}=2.29$ ,  $p=0.016$ ), the field of knowledge ( $F_{2, 415}=12.78$ ,  $p=0.001$ ) and the assessment satisfaction ( $F_{3, 415}=9.5196$ ,  $p=0.001$ ). No differences were found associated with teaching experience ( $F_{2, 415}=1.3466$ ,  $p=0.225$ ) or gender ( $F_{2, 415}=1.5486$ ,  $p=0.165$ ). To analyse possible significant differences in how teachers design assessment tasks using contextual characteristics such as the university, field of knowledge, experience or satisfaction with the assessment system, comparative analysis was performed, following a top-down strategy, using the permutational multivariate analysis of the variance (PERMANOVA). This began by analysing the differences using the four overall dimensions and continued until all the items were compared simultaneously. These analysis results are given in Table 5. Finally, the Kruskal-Wallis H test was applied to analyse each item individually, and the results are presented in Table 6.

The *post-hoc* analysis, using peer comparison, showed that significant differences between universities were found between UNIOVI and UCA ( $p=0.015$ ), and between UPV/EHU and UCA ( $p=0.002$ ). No significant differences were found in the remaining university pairings.

As for the field of knowledge, differences were found between EDU and ECO ( $p=0.001$ ), and between ECO and COM ( $p=0.003$ ). Finally, regarding the assessment security and satisfaction, differences appeared between Level 3 and Level 1 ( $p=0.022$ ), Level 1 and Level 2 ( $p=0.001$ ), Level 1 and Level 4 ( $p=0.001$ ) and between Level 2 and Level 4 ( $p=0.004$ ).

To analyse this in greater detail, further analysis (PERMANOVA II) took the dependent variables to be each of the items in

the self-report that made up each of the dimensions, making it possible to analyse the differences using a complete pattern of multiple dependent variables. Table 5 presents the results of this multivariate analysis, finding significant differences ( $p<.05$ ) among the four dimensions according to the university (UNI) and the assessment satisfaction (SSE). Differences due to the field of knowledge are only found in the RET and PAR dimensions. In the case of gender and experience, no significant differences were found.

A third analysis (PERMANOVA III) was performed from the multivariate perspective, considering all items individually as dependent variables. This analysis brought up significant differences ( $p<.05$ ) for the groups depending on the university, the field of knowledge and the assessment satisfaction.

Finally, each item was examined independently using the Kruskal-Wallis H test (Table 6). Significant differences ( $p<.05$ ) were found among the groups depending on the university in several items, such as I01, I04, I05 and I35 of the transparency dimension, in all the items related to the depth dimension, item I39 of the feedback dimension and I42 of the participation in the assessment dimension. Differences were also seen in the field of knowledge, such as item I05 of the transparency dimension, I41 of the feedback dimension and all the items from the participation in the assessment dimension. Furthermore, differences were found according to gender in item I42 of the participation in the assessment dimension, according to experience in items I03 and I05 of the transparency dimension, and according to satisfaction with the assessment in all the items being analysed, except for item I01.

### ***The university as the differentiating element***

The differences by university are specified in the TRA dimension in item I04, referring to offering information on the different participative assessment methods. In this item, the UPV/EHU scores ( $M=4.27$ ) are lower than for UNIOVI ( $M=4.81$ ), UCA ( $M=4.75$ ) or UDC ( $M=4.62$ ). In addition, in item I05,

alluding to providing information on the assessment tasks through guides, appendices, transparencies, etc., the significant difference appears between UPV/EHU (M=1.94) and UCA (M=2.81).

In the PRO dimension, the differences between universities are presented in the three items within this dimension. Consequently, in the item referring to using assessment tasks that imply using relevant subject knowledge and content (I36), the UPV/EHU presents the lowest scores (M=4.46) compared to UNIOVI (M=4.85) and UCA (M=4.87). In turn, referring to presenting students with challenging tasks (I37), the difference is seen between UPV/EHU (M=4.0) and UCA (M=4.51). Finally, regarding task authenticity (I38), meaning the assessment tasks that make it easier to transfer knowledge and skills into a professional context, the difference is seen between UPV/EHU (M=4.18) and UCA (M=4.59) and UDC (M=4.66).

The difference between universities in the case of the RET dimension revolves around the item regarding feedback that teachers give students (I39). Differences are determined between the URV (M=4.6) compared to the UDC (M=3.64) and the UPV/EHU (M=3.8), and also between the UV (M=4.3) and the UDC (M=3.64) and UPV/EHU (M=3.8).

Of the various participative activities considered in the dimension (PAR), the difference between universities is determined by student collaboration in specifying some elements of the assessment system (I42). This is the case for URV (M=1.5) which stands out from the other universities: UPV/EHU (M=2.62), UDC (M=2.64), UV (M=2.96) and UCA (M=3.21).

### ***Differences according to the field of knowledge***

By tackling possible differences that the teachers explain in their perceptions based on the field of knowledge they teach, significant differences are confirmed in the information, feedback and participation dimensions. In the case of information (I05), the Communication (M=2.72) and Education (M=2.40) teachers

stand out from Economics and Business Studies (M=1.99). This difference is also seen in I41 comparing the fields of Communication (M=3.5) and Education (M=2.89) against Economics and Business (M=2.10).

Finally, regarding participation, differences are found in the four items in this dimension. In the item alluding to student collaboration in specifying the assessment elements (I42), Communication (M=3.33) and Education (M=3.02) stand out from Economics and Business (M=2.11). In the self-assessment (I43), the difference is clear between Communication (M=2.67) and Education (M=3.02) compared to Economics and Business (M=1.46). We find a similar situation with peer assessment (I44) where Communication (M=2.67) and Education (M=2.45) stand out from Economics and Business (M=1.33) and in the use of dialogue-based and consensual assessment (I45), with higher scores in Education (M=1.81) and Communication (1.89) compared to Economics (M=1.06).

Table 5. PERMANOVA results depending on the university (UNI), field of knowledge (AMB), gender (GEN), experience (EXP) and assessment satisfaction (SSE)

Dimensions		UNI		AMB		GEN		EXP		SSE	
		F	Sig.	F	Sig.	F	Sig.	F	Sig.	F	Sig.
Overall dimensions	PERMANOVA I	2.2948	.016	12.78	.001	1.5486	.165	1.3466	.225	9.5196	.001
TRA Dimension Items		2.7512	.003	2.2213	.071	1.3037	.258	3,414	.009	4.6197	.001
PRO Dimension Items		2,780	.003	0.6418	.637	0.7603	.480	1.1895	.299	8.1473	.001
RET Dimension Items	PERMANOVA II	1.8399	.044	6.4295	.001	0.7483	.564	2.0355	.087	8.7211	.001
PAR Dimension Items		2.2845	.013	13,792	.001	1.8584	.094	0.5718	.705	3.9533	.001
RAPEVA items	PERMANOVA III	2.2938	.004	8.4893	.001	1.3697	.174	1.5244	0,126	5,657	.001

Table 6. Kruskal-Wallis H-test results depending on the university (UNI), field of knowledge (AM), gender (GEN), experience (EXP) and assessment satisfaction (SSE)

<i>Information</i>	UNI		AMB		GEN		EXP		SSE	
	H	Sig.	H	Sig.	H	Sig.	H	Sig.	H	Sig.
I01	11,064	.050	.407	.816	.513	.774	.968	.968	6,704	.082
I03	4,527	.476	.011	.995	.908	.635	6,758	.034	19,521	.000
I04	26,790	.000	1,294	.524	1,813	.404	5,776	.056	49,785	.000
I05	13,969	.016	6,396	.041	3,420	.181	9,449	.009	9,338	.025
I35	11,770	.038	3,725	.155	3,748	.154	.291	.865	21,021	.000
<i>Depth</i>										
I36	26,459	.000	1,593	.451	1,958	.376	1,087	.581	28,092	.000
I37	12,495	.029	.821	.663	2,311	.315	3,982	.137	28,751	.000
I38	16,715	.005	.790	.674	1,246	.536	2,183	.336	33,316	.000
<i>Feedback</i>										
I39	22,159	.000	5,435	.066	.970	.616	1,122	.571	23,642	.000
I40	6,756	.239	4,875	.087	1,520	.468	2,291	.318	44,559	.000
I41	7,138	.211	20,434	.000	.934	.627	5,091	.078	17,730	.000
<i>Participation</i>										
I42	19,585	.001	26,518	.000	7,645	.022	.065	.968	8,344	.039
I43	7,589	.180	24,533	.000	2,606	.272	3,828	.147	9,432	.024
I44	7,003	.220	31,239	.000	3,926	.140	1,337	.512	8,378	.039
I45	9,053	.107	20,538	.000	.280	.869	.876	.645	11,334	.010

### ***Teaching experience as a differentiating element***

In the case of experience, significant differences were only found between the teachers in the information dimension in items I03 and I05.

In item I03, teachers with over 20 years of experience (M=4.78), compared to teachers with less experience, between 11 and 20 years (M=4.66), consider that the assessment system informs students about what they must hand in or perform.

In addition, a difference is demonstrated between teachers with fewer years of experience (M=2.67) and those with more experience (M=2.02) regarding the information that is offered on the different participative methods of assessment such as self-assessment, peer assessment or co-assessment (I05).

### ***Assessment satisfaction as a differentiating element***

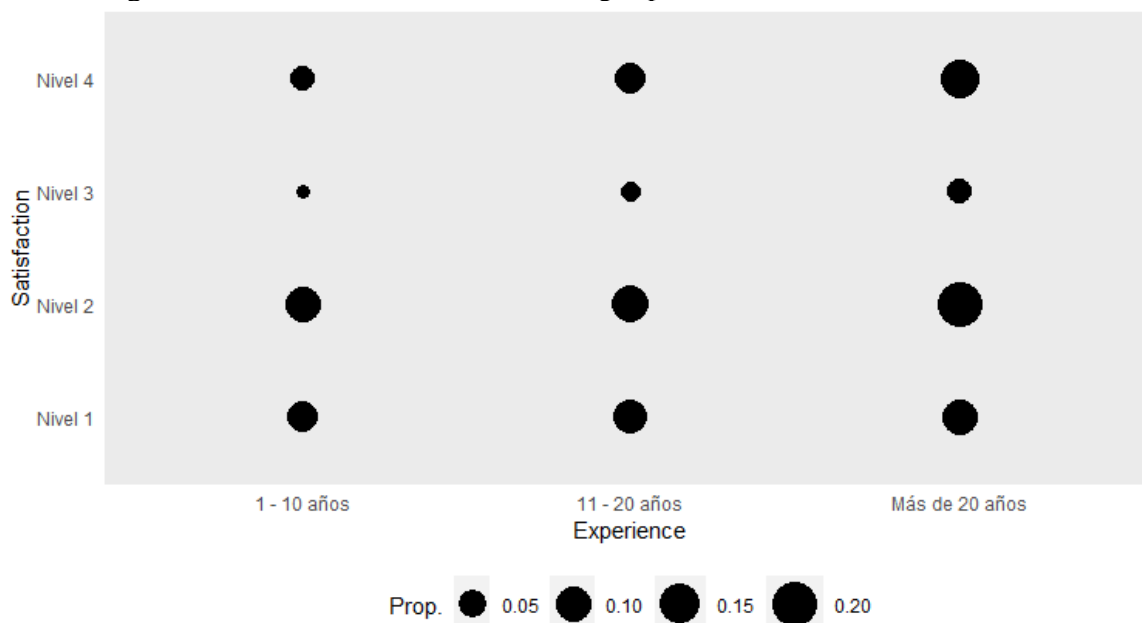
The assessment satisfaction levels have been built around the teachers' own

perceptions in their response to the indicators on their security regarding the assessment system (I48) and their degree of satisfaction with this system (I49). Working from the scores achieved in these two items, four levels were considered, taking the quartiles as a cut-off point.

The Chi-square test was performed to ascertain whether these satisfaction levels might be associated with teaching experience, obtaining a value of 6.149 (p=.407), thereby concluding that there was no association between these variables. In fact, teachers from each level of experience are spread over various levels of assessor satisfaction, as shown in the bubble diagram presented in Figure 4. For example, 22.1% of teachers with over 20 years of experience are in level 1, 42.6% in level 2, 7.8% in level 3 and 27.5% in level 4.

Based on this classification depending on the level of assessor satisfaction, significant differences are seen in all self-report items (Table 6) and, in all cases, the most obvious differences lie between level 1 and level 4, in favour of the latter.

Figure 4. Distribution of levels of teaching experience and evaluative satisfaction



## Discussion

This study firstly aims to analyse how university teachers characterise assessment tasks that they design in their assessment practice. Secondly, it looks for differences in this characterisation depending on their university, their field of knowledge, gender, experience or assessor satisfaction. The results of this work suggest implications, from both a theoretical and practical perspective and, in turn, outline some future lines of research.

### *Theoretical implications*

The initial contribution from this paper is its characterisation of assessment tasks into relevant four elements. The results obtained demonstrate that, from the teachers' perception, it is common practice to provide students with information on the assessment process, so that the students can amass all the information and guidance required to achieve the chosen learning outcomes. These results are consistent with contributions from various authors such as Biggs & Tang (2011) and Boud (2022) who consider that students need to be clearly informed on what they have to do and the minimum standards that must be reached, although it should be highlighted that students lack information relating to the participative assessment methods, particularly from teachers with more experience. In addition, it has highlighted that teachers believe they design assessment tasks in great depth, making it easier for students to achieve deeper learning.

However, it shows that teacher still resist using feedback or, to an even greater extent, participation from students in the evaluative process, despite the clear benefits implied for students, as demonstrated by papers from Panadero et al. (2017), Tai et al. (2018) and reviews by Carless & Boud (2018) and Carless & Winstone (2020).

A second implication refers to the importance of contextual aspects. In this respect, the results demonstrate that the university where the teachers work, or their specific field of knowledge, are possible

elements that modulate evaluative practice. However, satisfaction with the assessment is the variable which differs the most among teachers. This might be because the teachers consider that their assessment activity has been designed in line with their grounding, knowledge and experience. However, this must be questioned because, as mentioned by DeLuca & Johnson (2017), many teachers make decisions on assessment without enough experience or skills training and, as stated by Mehrabi & Hosseini (2021), without a solid model as a guide for teachers to use to design assessment to solve problems together.

### *Practical implications*

From a practical perspective, based on the outcomes, this requires boosting and developing training for university teachers, focusing on the critical attitude required to encourage the necessary changes, because as stated by Lo & Leung (2022) to develop an effective educational assessment, it is essential to work on assessment literacy among teachers. The technology boost, due to the Covid-19 pandemic, only emphasises the need for this critical attitude in the light of modernisation and changes. Technology can ease feedback processes and student participation in assessment processes (Gómez-Ruiz et al., 2020; Ibarra-Sáiz & Rodríguez-Gómez, 2017; Rodríguez-Gómez & Ibarra-Sáiz, 2016) which would improve students' reflection and self-regulation, but their learning can only improve if they use the information they are given appropriately and effectively.

### *Limitations and future research*

Methodologically, this paper presents a series of limitations that, in turn, constitute new lines of future research. Firstly, the study has been carried out in a specific context in Spanish public universities and it should be extended to other public and private universities, and to other countries to provide a comparative analysis. Secondly, this study has worked from the teachers' perceptions. Consequently, extending assessment practice through in-depth analysis using a multiple case

study would provide in-depth analysis of the type of assessment tasks designed by the teaching staff and how they are perceived and experienced by the students.

In addition, it would be interesting to examine the teachers' satisfaction with their assessment practice and their criticism of their own actions, because as stated by Boud (2020, p. 7) "An outsider would be astounded to discover how much practice still occurs which cannot be defended on the basis of any scholarship of assessment". In this way, better comprehension can be achieved not only on the assessment practice but above all regarding the conceptions that determine these practices to a large extent, to the point that there is a major lack of connection between the scores awarded to the students and the expected learning outcomes for a course or subject (Boud, 2020).

## Conclusion

The student's learning outcomes can be strengthened if the quality of assessment tasks designed is good enough. This means that the tasks should be transparent, they should reinforce in-depth learning, encourage feedback and make students part of the assessment process.

This study has analysed the perception of teachers from different universities regarding the tasks that they design. It has been seen that the transparency is generally appropriate. This refers to the information given to the student on what they are going to be asked to do to achieve the intended learning outcomes. In the same way, the teachers demonstrate that they require in-depth knowledge from their students regarding which complex relationships must be determined between fundamental concepts by implicating them in tasks that are contextualised in their future professional practice.

On the contrary, the outlook is not as bright regarding feedback and participation. The teachers mention that they provide students with feedback on their progress in the teaching-learning process. However, participation from the students is still very low

in this feedback, either self-assessed or peer assessed. This is also the position regarding student participation in assessment tasks, given that they are not offered the chance to collaborate in specifying some elements of the assessment system; nor is the assessment or scoring strengthened in a negotiated, consensual way between teachers and students.

Neither gender nor teaching experience are associated with the teachers' perception of their evaluative practice. However, there are differences between teachers from different universities and fields of knowledge. In the latter case, teachers on the master's degrees in economics and business studies mention less practice than those teaching education and communication regarding the use of feedback or student participation in the assessment. Finally, the teachers' degree of satisfaction with their own assessment practice is shown to be determining. The higher the level of satisfaction, the better their perception of assessment task quality. In any case, it has been seen that teacher training must be boosted to promote the use of feedback and, above all, make it easier for students to take part in the assessment, which would improve the quality of the assessment tasks and the chosen learning outcomes.

## Funding & Acknowledgments

This work was made possible by the FLOASS Project (Ref. RTI2018-093630-B-I00) funded by Spanish Ministry of Science, Innovation and Universities (MCIU), State Research Agency (AEI) and the European Regional Development Fund (ERDF), and support from the UNESCO Chair in Assessment & Evaluation, Innovation and Excellence in Education at the University of Cadiz.

## References

- Anderson, M. J. (2017). Permutational Multivariate Analysis of Variance (PERMANOVA). *Wiley StatsRef: Statistics Reference Online*, 1–15. <https://doi.org/10.1002/9781118445112.sta>

- [t07841](#)
- Bearman, M., Dawson, P., Boud, D., Bennett, S., Hall, M., & Molloy, E. (2016). Support for assessment practice: developing the Assessment Design Decisions Framework. *Teaching in Higher Education*, 21(5), 545–556. <https://doi.org/10.1080/13562517.2016.1160217>
- Bearman, M., Dawson, P., Boud, D., Hall, M., Bennett, S., Molloy, E., & Joughin, G. (2014). *Guide to the assessment design decisions framework*. <http://www.assessmentdecisions.org/guide/>
- Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university. What the students does* (4th ed.). McGraw-Hill-SRHE & Open University Press.
- Boud, D. (2020). Challenges in reforming higher education assessment: a perspective from afar. *RELIEVE*, 26(1), art. M3. <https://doi.org/10.7203/relieve.26.1.17088>
- Boud, D. (2022). Assessment-as-learning for the development of students' evaluative judgement. In Z. Yan & L. Yang (Eds.), *Assessment as Learning. Maximising Opportunities for Student Learning and Achievement* (pp. 25–37). Routledge. <https://doi.org/10.4324/9781003052081-3>
- Boud, D., & Dawson, P. (2021). What feedback literate teachers do: an empirically-derived competency framework. *Assessment & Evaluation in Higher Education*, April, 1–14. <https://doi.org/10.1080/02602938.2021.1910928>
- Carless, D. (2015). Exploring learning-oriented assessment processes. *Higher Education*, 69(6), 963–976. <https://doi.org/10.1007/s10734-014-9816-z>
- Carless, D. (2022). From teacher transmission of information to student feedback literacy: Activating the learner role in feedback processes. *Active Learning in Higher Education*, 23(2), 143–153. <https://doi.org/10.1177/1469787420945845>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Carless, D., & Winstone, N. (2020). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education*, 1–14. <https://doi.org/10.1080/13562517.2020.1782372>
- DeLuca, C., & Johnson, S. (2017). Developing assessment capable teachers in this age of accountability. *Assessment in Education: Principles, Policy and Practice*, 24(2), 121–126. <https://doi.org/10.1080/0969594X.2017.1297010>
- Falchikov, N. (2005). *Improving assessment through student involvement. Practical solutions for aiding learning in higher education and further education*. RoutledgeFalmer.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A Meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287–322. <https://doi.org/10.3102/00346543070003287>
- Gómez-Ruiz, M. Á., Ibarra-Sáiz, M. S., & Rodríguez-Gómez, G. (2020). Aprender a evaluar mediante juegos de simulación en educación superior: percepciones y posibilidades de transferencia para los estudiantes. *Revista Iberoamericana de Evaluación Educativa*, 13(1), 157–181. <https://doi.org/10.15366/rie2020.13.1.007>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial Least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage. <https://doi.org/10.1007/978-3-030-80519-7>
- Henderson, M., Boud, D., Molloy, E., Dawson, P., Phillips, M., Ryan, T., & Mahoney, P. (2018). *Feedback for learning: Closing the assessment loop*. Department of Education and Training. <https://nla.gov.au/nla.obj-719788718/view>



- Henseler, J. (2021). *Composite-based structural equation modeling. Analyzing latent and emergent variables*. Guilford Press.
- Hortigüela Alcalá, D., Palacios Picos, D., & López Pastor, V. (2019). The impact of formative and shared or co-assessment on the acquisition of transversal competences in higher education. *Assessment & Evaluation in Higher Education*, 44(6), 933–945. <https://doi.org/10.1080/02602938.2018.1530341>
- Ibarra-Sáiz, M. S., & Rodríguez-Gómez, G. (2020). Evaluando la evaluación. Validación mediante PLS-SEM de la escala ATAE para el análisis de las tareas de evaluación. *RELIEVE*, 26(1), art. M4. <https://doi.org/10.7203/relieve.26.1.17403>
- Ibarra-Sáiz, M.S., & Rodríguez-Gómez, G. (2017). EvalCOMIX®: A web-based programme to support collaboration in assessment. In T. Issa, P. Kommers, T. Issa, P. Isaías, & T. B. Issa (Eds.), *Smart technology applications in business environments* (pp. 249–275). IGI Global. <https://doi.org/10.4018/978-1-5225-2492-2.ch012>
- Ibarra-Sáiz, M.S., Rodríguez-Gómez, G., & Boud, D. (2021). The quality of assessment tasks as a determinant of learning. *Assessment & Evaluation in Higher Education*, 46(6), 943–955. <https://doi.org/10.1080/02602938.2020.1828268>
- Ibarra-Sáiz, M.S., Rodríguez-Gómez, G., Lukas-Mujika, J.F., & Santos-Berrondo, A. (2023). Medios e instrumentos para evaluar los resultados de aprendizaje en másteres universitarios. Análisis de la percepción del profesorado sobre su práctica evaluativa. *Educación XXI*, 26(1), 21–45. <https://doi.org/10.5944/educxx1.33443>
- JASP Team. (2022). *JASP (Version 0.16.1)*. <https://jasp-stats.org/>
- Johnson, R. L., & Morgan, G. B. (2016). *Survey scales. A guide to development, analysis, and reporting*. The Guilford Press.
- Lipnevich, A. A., & Panadero, E. (2021). A Review of Feedback Models and Theories: Descriptions, Definitions, and Conclusions. *Frontiers in Education*, 6(December). <https://doi.org/10.3389/feduc.2021.720195>
- Lo, Y. Y., & Leung, C. (2022). Conceptualising assessment literacy of teachers in content and language integrated learning programmes. *International Journal of Bilingual Education and Bilingualism*, 1–19. <https://doi.org/10.1080/13670050.2022.2085028>
- López Gil, M., Gómez Ruiz, M. Á., Vázquez Recio, R., & Ruiz Romero, A. (2022). La pesadilla de la evaluación: Análisis de los sueños de estudiantes universitarios. *Revista Iberoamericana de Evaluación Educativa*, 15(1), 139–159. <https://doi.org/10.15366/rie2022.15.1.008>
- Mehrabi Boshrabadi, A., & Hosseini, M. R. (2021). Designing collaborative problem solving assessment tasks in engineering: an evaluative judgement perspective. *Assessment & Evaluation in Higher Education*, 46(6), 913–927. <https://doi.org/10.1080/02602938.2020.1836122>
- Panadero, E., & Alqassab, M. (2019). An empirical review of anonymity effects in peer assessment, peer feedback, peer review, peer evaluation and peer grading. *Assessment & Evaluation in Higher Education*, 44(8), 1253–1278. <https://doi.org/10.1080/02602938.2019.1600186>
- Panadero, E., Fraile, J., & García-Pérez, D. (2022). Transición a educación superior y evaluación: un estudio longitudinal anual. *Educación XXI*, 25(2), 15–37. <https://revistas.uned.es/index.php/educacionXXI/article/view/29870>
- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74–98. <https://doi.org/10.1016/j.edurev.2017.08.004>

- Quesada-Serra, V., Gómez Ruiz, M. A., Gallego Noche, M. B., & Cubero-Ibáñez, J. (2019). Should I use co-assessment in higher education? Pros and cons from teachers and students' perspectives. *Assessment & Evaluation in Higher Education*, 44(7), 987-1002. <https://doi.org/10.1080/02602938.2018.1531970>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Rodríguez-Gómez, G., & Ibarra-Sáiz, M. S. (2015). Assessment as learning and empowerment: Towards sustainable learning in higher education. In M. Peris-Ortiz & J. M. Merigó Lindahl (Eds.), *Sustainable learning in higher education. Developing competencies for the global marketplace* (pp. 1–20). Springer International Publishing. [https://doi.org/10.1007/978-3-319-10804-9\\_1](https://doi.org/10.1007/978-3-319-10804-9_1)
- Rodríguez-Gómez, G., & Ibarra-Sáiz, M. S. (2016). Towards sustainable assessment: ICT as a facilitator of self- and peer assessment. In M. Peris-Ortiz, J. A. Gómez, F. Vélez-Torres, & C. Rueda-Armengot (Eds.), *Education tools for entrepreneurship* (pp. 55–71). Springer International Publishing. [https://doi.org/10.1007/978-3-319-24657-4\\_5](https://doi.org/10.1007/978-3-319-24657-4_5)
- Sambell, K., McDowell, L., & Montgomery, C. (2013). *Assessment for Learning in Higher Education*. Routledge. <https://doi.org/10.4324/9780203818268>
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: enabling students to make decisions about the quality of work. *Higher Education*, 76(3), 467–481. <https://doi.org/10.1007/s10734-017-0220-3>
- Yan, Z., & Boud, D. (2022). Conceptualising assessment-as-learning. In Z. Yan & L. Yang (Eds.), *Assessment as learning. Maximising opportunities for student learning and achievement* (pp. 11–24). Routledge. <https://doi.org/10.4324/9781003052081-2>

#### Authors / Autores

**Ibarra-Sáiz, M<sup>a</sup>. Soledad** ([marisol.ibarra@uca.es](mailto:marisol.ibarra@uca.es))  0000-0003-4513-702X

Senior Lecturer in the Area of Research, Evaluation and Assessment Methods in Education at University of Cadiz. Director of UNESCO Chair on Evaluation and Assessment, Innovation and Excellence in Education. Director of EVALfor Research Group-SEJ509 “Assessment & Evaluation in Training Contexts”. She develops her research mainly in the field of assessment and evaluation in higher education. She has been principal researcher of more than ten international and national projects, whose results have been published in various articles, book chapters and contributions to international conferences. She is currently the main co-researcher of the ValeFLOASS Project - Valorisation of the e-FLOASS environment for the monitoring and assessment of learning outcomes in the educational system (PDC2022—133050-I00) in which nine Spanish universities participate.

**Lukas-Mujika, José-Francisco** ([jf.lukas@ehu.eus](mailto:jf.lukas@ehu.eus))  0000-0002-9440-035X

Senior Lecturer in Educational Assessment and Evaluation at University of the Basque Country. Faculty of Education, Philosophy and Anthropology. Department of Educational Sciences. Expert in educational measurement and evaluation with a special focus on issues related to the evaluation of educational programs and institutions and the construction of measurement instruments. Deputy Director of the Institute of Educational Sciences of the University of the Basque Country (1997-2000). Director of the Department of Research and Diagnostic Methods in Education of the University of the Basque Country (2003-2009). Member of the Scientific Committee of the Basque Institute for Non-University Educational Evaluation and Research (ISEI-IVEI) (2002-2014). Evaluator in different evaluation and accreditation agencies (ANECA, DEVA, ACSUCYL). Author of articles, books, book chapters and contributions to international conferences.

**Ponce-González, Nicolás** ([nicolas.ponce@uca.es](mailto:nicolas.ponce@uca.es))  0000-0003-1336-6303

PhD in Education Sciences from the University of Granada. Interim Professor in the Area of Research, Evaluation and Assessment Methods in Education at the University of Cadiz. He is part of the team of the UNESCO Chair on Evaluation and Assessment, Innovation and Excellence in Education. He has participated in research projects at the regional, national, and international levels. He is currently a member of the research team of the ValeFLOASS Project - Valorisation of the e-FLOASS environment for monitoring and assessment of learning outcomes in the educational system (PDC2022-133050-I00093630-B-I00). He is also a member of EVALfor Research Group-SEJ509 "Assessment & Evaluation in Training Contexts". His main research lines are training for the development of competencies, evaluation of intervention programs in Emotional Intelligence, and evaluation in Higher Education. He is a co-author of several publications of articles, book chapters, and communications to national and international congresses.

**Rodríguez-Gómez, Gregorio** ([gregorio.rodriguez@uca.es](mailto:gregorio.rodriguez@uca.es))  0000-0001-9337-1270

Professor of Educational Research Methods at the University of Cadiz. He is the coordinator of the strategic area "Studies and research in assessment and evaluation" of the UNESCO Chair on Evaluation and Assessment, Innovation and Excellence in Education. Founding member of the EVALfor Research Group "SEJ509" Assessment and evaluation in training contexts. His research interest is focused on research methods and assessment and evaluation in higher education. He is currently the main co-researcher of the ValeFLOASS Project - Valorisation of the e-FLOASS environment for the monitoring and assessment of learning outcomes in the educational system (PDC2022-133050-I00) in which nine Spanish universities participate. Author of articles, book chapters and contributions to international conferences. He has been President of the Interuniversity Association for Pedagogical Research (AIDIPE) and the Spanish University Teaching Network (RED-U.)



**Revista ELectrónica de Investigación y EValuación Educativa**  
*E-Journal of Educational Research, Assessment and Evaluation*

[ISSN: 1134-4032]



Esta obra tiene [licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/).

This work is under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).