

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

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

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看法

تصور أعضاء هيئة التدريس بالجامعة لجودة مهام تقييم مخرجات التعلم

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

⁽¹⁾ Cátedra UNESCO en Evaluación, Innovación y Excelencia en Educación, Universidad de Cádiz, España.

⁽²⁾ Universidad del País Vasco /Euskal Herriko Unibertsitatea, España

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.

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.

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 المواعمة 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

Introducción

En el diseño de los procesos de evaluación son numerosas las decisiones que debe tomar el profesorado sobre una gran cantidad de elementos curriculares (Bearman et al., 2014, 2016). Este trabajo se centra en la calidad de las tareas de evaluación ya que constituye un elemento esencial del proceso evaluativo. 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.

En esta aportación se explora y analiza la práctica evaluativa del profesorado universitario, a partir de sus propias percepciones y experiencia, centrando la

atención específicamente en la naturaleza de las tareas de evaluación que diseña, es decir, las características que pueden servir como elementos diferenciadores de unas tareas de evaluación frente a otras.

Desde diferentes enfoques evaluativos, los trabajos de Carless (2015), Rodríguez-Gómez e Ibarra-Sáiz (2015) o Sambell et al. (2013) destacan la relevancia del diseño de las tareas de evaluación. Pero ¿cómo caracteriza el profesorado universitario las tareas de evaluación que diseña en su práctica evaluativa?, ¿existe alguna diferencia en el diseño de estas tareas de evaluación que pueda asociarse a características como la universidad de pertenencia, el ámbito de conocimiento, el género, los años de experiencia o la seguridad y satisfacción evaluadora percibida?

Para dar respuesta a estos interrogantes, sobre la base de la percepción del profesorado, en esta investigación se pretende:

- analizar cuáles son las características de las tareas de evaluación que diseña el profesorado universitario;
- comprobar si se pueden asociar diferencias en el diseño de las tareas de evaluación a partir de ciertas características contextuales como la universidad, el ámbito de conocimiento, la experiencia o la satisfacción con el sistema de evaluación.

Las tareas de evaluación y sus características

La evaluación exige dar respuesta a dos cuestiones básicas: ¿para qué evaluar? y ¿qué evaluar? A partir de la respuesta que se ofrezca a estas dos cuestiones nos encontramos con diferentes enfoques evaluativos en la educación superior. Sobre la base de las aportaciones de autores como Boud (2022), Ibarra-Sáiz et al. (2021) y Sambell et al. (2013) en la Tabla 1 se presenta una breve síntesis de los tres enfoques fundamentales.

Tabla 1. Enfoques sobre la evaluación y el aprendizaje

Enfoque evaluativo	¿Para qué evaluar?	¿Qué evaluar?
Evaluación <i>del</i> aprendizaje	Para certificar o informar al estudiante y otros interesados de la competencia del estudiante en relación con los resultados de aprendizaje	La medida en que los estudiantes pueden aplicar los conceptos clave, el conocimiento, las habilidades y las actitudes relacionadas con los resultados de aprendizaje
Evaluación <i>para el</i> aprendizaje	Para permitir al profesorado que determine cuáles son los siguientes pasos para avanzar en el aprendizaje del estudiante	Cada progreso del estudiante y las necesidades de aprendizaje en relación con los resultados de aprendizaje deseados
Evaluación <i>como</i> aprendizaje	Para guiar y aportar oportunidades para que cada estudiante pueda monitorizar y reflexionar críticamente sobre su aprendizaje e identificar los siguientes pasos a seguir	El pensamiento de cada estudiante sobre su aprendizaje, qué estrategias utiliza para apoyar o retar ese aprendizaje y los mecanismos que utiliza para adecuar y avanzar en su aprendizaje

Desde cualquiera de estos enfoques, las tareas de evaluación juegan un papel central. En primer lugar, la naturaleza de las tareas de evaluación se relaciona con el enfoque que los estudiantes adoptan en su aprendizaje (Biggs & Tang, 2011). En segundo lugar, porque como han puesto de manifiesto los trabajos de Panadero et al. (2022) y López Gil et al. (2022), las actividades de evaluación tiene un gran efecto sobre las emociones y la

motivación de los estudiantes. Y, por último, las tareas de evaluación constituyen la base de las evidencias sobre las cuales el profesorado podrá valorar el grado de consecución de los resultados de aprendizaje, las posibles modificaciones necesarias para mejorar el proceso de enseñanza-aprendizaje, u orientar al estudiante sobre las estrategias a utilizar para impulsar su progreso y avanzar en su aprendizaje.

Este estudio se centra en cuatro elementos esenciales de la naturaleza de las tareas: transparencia, profundidad, retroalimentación y participación. La inclusión de estos elementos en el diseño de las tareas de evaluación favorecerá que éstas sean creíbles, rigurosas, interesantes y promovedoras de un aprendizaje valioso para el estudiante, es decir, tareas de evaluación de calidad (Ibarra-Sáiz et al., 2021).

La transparencia se refiere a la necesaria información que ha de facilitarse al estudiante, lo que Gore et al. (2009) conceptualizaron como apoyo al estudiante. Así, la tarea de evaluación ha de ser pública, conocida e interiorizada por el estudiante, es preciso que el estudiante reconozca la calidad de los productos o actuaciones que tendrá que realizar para alcanzar los resultados de aprendizaje esperados; sepa qué es lo que se espera que aprenda y ejecute en la tarea; identifique los medios e instrumentos de evaluación que se van a utilizar para valorar su desempeño; y saber cuál va a ser su papel como evaluador a través de modalidades como la autoevaluación, evaluación entre iguales y/o coevaluación. Como destacan Yan y Boud (2022), para poder desarrollar el aprendizaje de los estudiantes las tareas de evaluación han de permitir que el juicio sobre los resultados de aprendizaje sea realizado no sólo por el profesorado, sino también por los propios estudiantes. Y para ello será preciso ofrecer a los estudiantes la información y orientaciones básicas que le faciliten abordar las tareas de forma exitosa.

Una tarea de evaluación de calidad exige que el estudiante establezca relaciones complejas y coherentes entre conceptos fundamentales y significativos, lo que implica un conocimiento profundo (Gore et al., 2009). Asimismo, uno de los cambios importantes que han tenido lugar en el último decenio ha sido el aumento de la evaluación auténtica (Boud, 2020) mediante el uso de tareas complejas (Sambell et al., 2013), lo que implica utilizar tareas y procesos contextualizados en la práctica profesional. Este tipo de tareas no sólo mejoran la motivación y el compromiso de los

estudiantes sino que ofrecen oportunidades para que los estudiantes resuelvan problemas de la vida real y son más significativas para su aprendizaje (Yan & Boud, 2022). En definitiva, tareas que promuevan el aprendizaje profundo a través del uso de métodos de indagación y el uso del pensamiento crítico y reflexivo (Ibarra-Sáiz & Rodríguez-Gómez, 2020).

Las aportaciones de Henderson et al. (2018), Carless (2020), Lipnevich & Panadero (2021) y Boud & Dawson (2021) evidencian la importancia y relevancia de la retroalimentación en el proceso de evaluación. En consecuencia, el diseño de las tareas de evaluación ha de prever cómo y de qué forma se utilizará y promoverá esta retroalimentación, el papel que jugarán los estudiantes en ofrecer y recibir retroalimentación, y el uso que harán los estudiantes de la información facilitada.

La participación del estudiante en el proceso de evaluación puede materializarse en diferentes momentos y situaciones. Desde el diseño hasta la calificación, desde la redacción de los criterios de evaluación al diseño y concreción de los medios e instrumentos de evaluación. Los trabajos de Falchikov (2005), Falchikov & Goldfinch (2000) y las últimas aportaciones de Panadero & Alqassab (2019), Hortigüela et al. (2019) y Quesada-Serra et al. (2019) constatan el efecto positivo de la participación de los estudiantes en su proceso de aprendizaje, gracias al uso de modalidades evaluativas como la autoevaluación, la evaluación entre iguales o la coevaluación.

Método

Diseño de investigación

Se ha optado por una metodología mixta (qual->QUAN) mediante un diseño secuencial exploratorio (Creswell, 2015). En la primera fase se procedió al diseño y validación de contenido del cuestionario RAPEVA-Autoinforme del profesorado sobre su práctica en la evaluación de los resultados de aprendizaje. En la segunda fase, en el período final del curso académico 2020/21, se

recabaron las percepciones del profesorado de ciencias sociales pertenecientes a seis universidades.

El autoinforme RAPEVA

La construcción del autoinforme RAPEVA se inició con una revisión de la literatura y, posteriormente, se utilizó el método de consenso grupal (Johnson & Morgan, 2016) para la validación de contenido, en la que participaron 22 jueces en tres iteraciones. La definición y concreción de los diferentes indicadores se revisaron al finalizar cada una de las iteraciones.

El autoinforme RAPEVA está basado en combinaciones lineales de las variables observadas (Henseler, 2021). En la valoración que realiza cada encuestado sobre cada uno de los ítems del autoinforme juega un papel esencial los aspectos cognitivos y actitudinales, por lo que éstos ítems se constituyen en un índice formativo (Hair et al.,

2022). En consecuencia, se decidió realizar un análisis generalizado de componentes estructurales (Hwang & Takane, 2015), obteniéndose medidas de ajuste (GFI=.89 y SRMR=.08) que se consideran aceptables.

Respecto a su estructura, en la primera parte del autoinforme se solicita información sobre aspectos contextuales como la universidad de procedencia, el ámbito de conocimiento, años de experiencia o el género. En la segunda se presentan 49 ítems en formato de escala tipo Likert (0-5) estructurados en once dimensiones (Tabla 2). La cumplimentación del autoinforme requería unos 20 minutos.

Como se ha señalado anteriormente, este estudio presenta los resultados de las cuatro dimensiones relacionadas con la calidad de las tareas de evaluación (TRA, PRO, RET y PAR). Los resultados relacionados con los medios e instrumentos de evaluación pueden consultarse en el trabajo de Ibarra-Sáiz et al. (2023).

Tabla 2. Estructura del autoinforme RAPEVA

Dimensiones		# Ítems	Ítems
TRA *	Transparencia	5	I01, I03 al I05, I35
CAE	Competencias a evaluar	6	I06 al I11
MOB	Medios de observación	6	I12 al I17
MEN	Medios de encuesta	4	I18 al I21
MDA	Documentos y artefactos	9	I22 al I30
INE	Instrumentos de evaluación	4	I31 al I34
PRO*	Profundidad de las tareas	4	I36 al I38
RET*	Retroalimentación	3	I39 al I41
PAR*	Participación	4	I42 al I45
FOR	Formación en evaluación	2	I46, I47
SSE	Satisfacción con evaluación	2	I48, I49

Participantes

El cuestionario se envió en formato online a todo el profesorado de los másteres de Comunicación, Educación y Economía y Empresa de las seis universidades participantes en el Proyecto FLOASS (<http://floass.uca.es>): Universidad de Cádiz (UCA), Universidad Rovira i Virgili (URV), Universidad de Oviedo (UNIOVI), Universidad de Valencia (UV),

Universidad de A Coruña (UDC) y Universidad del País Vasco (UPV/EHU). Se invitó a todo el profesorado mediante un correo electrónico enviado por parte de las coordinaciones de los títulos, en el que se indicaba el enlace a través del cual cumplimentar el autoinforme. Sobre una muestra de 2.400 profesores invitados, iniciaron la cumplimentación un total de 626 profesores y se obtuvieron 416 autoinformes

completos de profesores que impartían docencia en 63 másteres diferentes, alcanzándose así un nivel de confianza del 95%. En la Tabla 3 se presenta la distribución de los 416 profesores en función de la

universidad de procedencia, el género, ámbito de conocimiento, años de experiencia y el grado de seguridad y satisfacción con la evaluación.

Tabla 3. Características demográficas

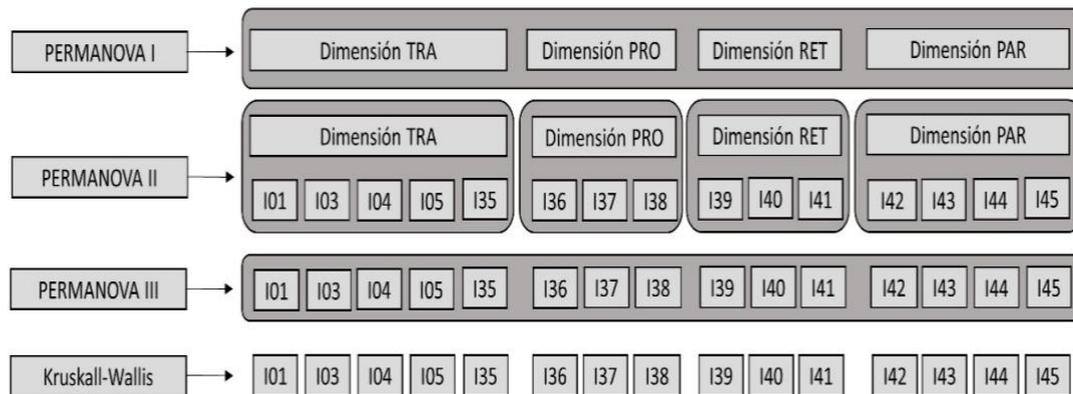
	Mujer		Hombre		Otros		Total	
	n	%	n	%	n	%	n	%
<i>Universidad</i>								
UCA	30	49.2	31	50.8	0	0	61	14.7
URV	18	60	12	40	0	0	30	7.2
UNIOVI	35	63.6	20	36.4	0	0	55	13.2
UV	36	51.4	33	47.1	1	1.4	70	16.8
UDC	45	54.2	37	44.6	1	1.2	83	20
UPV/EHU	61	52.1	56	47.9	0	0	117	28.1
<i>Ámbito</i>								
COM	10	55.6	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>Experiencia</i>								
<10 años	59	64.8	32	35.2	0	0	91	21.9
11-20 años	78	64.5	41	33.9	2	1.7	121	29.1
>20 años	88	43.1	116	56.9	0	0	204	49
<i>Satisfacción</i>								
Nivel 1	75	68.2	34	30.9	1	0.9	110	26.4
Nivel 2	104	59.1	72	40.9	0	0	176	42.3
Nivel 3	8	26.7	22	73.3	0	0	30	7.2
Nivel 4	38	38	61	61	1	1	100	24
Total	225	54.1	189	45.4	2	0.5	416	100

Análisis de datos

Para alcanzar el primer objetivo se ha realizado un análisis estadístico exploratorio y, para el segundo, se han analizado las diferencias entre grupos. Al tratarse de mediciones ordinales que no se ajustaban a la normalidad (Prueba de K-S, $p < .001$), se ha recurrido a diferentes técnicas no paramétricas. El PERMANOVA-Permutational Multivariate Analysis of Variance (Anderson, 2017) se ha utilizado en tres momentos diferentes (Figura 1). Esta técnica tiene un objetivo similar al MANOVA, por lo que se la conoce también como MANOVA no paramétrico, pero se diferencia en que se basa en el análisis de permutaciones sobre las matrices de distancias

para realizar la comparación multivariante. En primer lugar, se contrastaron las posibles diferencias entre grupos comparando las cuatro dimensiones globales (PERMANOVA I). En segundo lugar, con el objetivo de obtener una mayor comprensión de las diferencias encontradas, se compararon individualmente los ítems constitutivos de cada una de las dimensiones (PERMANOVA II). Por último, para profundizar en las relaciones entre variables se realizó un análisis multivariante con todos los ítems simultáneamente (PERMANOVA III). Finalmente se procedió al análisis individual de cada ítem mediante la prueba H de Kruskal-Wallis. Los análisis se han realizado con JASP (JASP Team, 2022) y R (R Core Team, 2021).

Figura 1. Proceso del análisis comparativo



Resultados

Inicialmente se presentan los resultados globales obtenidos en las respuestas del profesorado al autoinforme, organizados los mismos en las cuatro características de las tareas de evaluación analizadas: transparencia, retroalimentación, profundidad y participación. Posteriormente se expondrán los resultados sobre las diferencias encontradas en las percepciones tomando como variables de comparación la universidad, el ámbito de conocimiento, el género, la experiencia y la seguridad y satisfacción con la evaluación.

Caracterización de las tareas de evaluación

En la Tabla 4 se muestran las medidas de tendencia central (mediana y media) y dispersión (desviación típica) para cada uno de los ítems del autoinforme agrupados en las cuatro dimensiones referidas, la cual nos permite analizar las características de las tareas de evaluación diseñadas por el profesorado

universitario. Para la dimensión transparencia se presenta un mayor grado de acuerdo con afirmar que los resultados de aprendizaje, criterios, procedimientos de evaluación y sistemas de calificación de las asignaturas son públicos, apareciendo estos en las guías docentes/programas (I01, $M=4.71$).

En la dimensión profundidad el mayor grado de acuerdo se presenta al afirmar que las tareas de evaluación implican el uso de conocimiento y contenidos relevantes de la materia por el estudiante (I36, $M=4.69$). En relación con la dimensión retroalimentación el mayor grado de acuerdo está relacionado con proporcionar retroalimentación a los estudiantes sobre su progreso durante el proceso de enseñanza-aprendizaje (I39, $M=3.98$). En la dimensión participación las medias son muy bajas. El mayor grado de acuerdo, aun siendo bajo, se encuentra a la hora de facilitar que los estudiantes colaboren con la concreción de algunos elementos del sistema de evaluación (I42, $M=2.67$).

Tabla 4. Medidas de tendencia central y dispersión en las dimensiones e ítems del autoinforme RAPEVA

	Mdn	M	SD
Transparencia (TRA)	4.2	4.18	.58
I01. Los resultados de aprendizaje, criterios, procedimientos de evaluación y sistema de calificación de la asignatura son públicos (aparecen en la guía docente/programa).	5	4.71	.73
I03. En el sistema de evaluación se informa acerca de lo que los estudiantes tienen que entregar o ejecutar.	5	4.70	.66
I04. En el sistema de evaluación se informa acerca de los criterios e instrumentos de evaluación.	5	4.53	.81
I05. En el sistema de evaluación se informa de las cuatro modalidades de evaluación (autoevaluación, evaluación entre iguales, coevaluación/evaluación compartida y evaluación por el profesorado).	2	2.25	1.83
I35. Informo y describo a los estudiantes las tareas de evaluación que tienen que realizar (a través de guías, adendas, transparencias, etc.).	5	4.70	.71
Profundidad (PRO)	4.67	4.44	.71
I36. Las tareas de evaluación implican el uso de conocimientos y contenidos relevantes de la materia por el estudiante.	5	4.69	.68
I37. Las tareas de evaluación se plantean como un reto para el estudiante.	5	4.17	1.12
I38. Las tareas de evaluación facilitan la aplicación de conocimientos y habilidades a situaciones o casos parecidos a los que encontrará el estudiante en su futuro profesional.	5	4.45	.91
Retroalimentación (RET)	3.33	3.30	1.31
I39. Proporciono retroalimentación a los estudiantes sobre su progreso durante el proceso de enseñanza-aprendizaje (mediante la realización de borradores ofreciendo una revisión individual o sesiones/talleres sobre ellos, comparación con buenas prácticas o productos, etc.).	4	3.98	1.26
I40. Facilito que los estudiantes participen y se auto retroalimenten contrastando sus avances con los instrumentos de evaluación, comparación con buenas prácticas o productos, etc., que les he proporcionado previamente.	4	3.32	1.63
I41. Los estudiantes reciben retroalimentación sobre su progreso por parte de sus compañeros mediante la revisión entre iguales de borradores, presentaciones orales, etc.	3	2.60	1.91
Participación (PAR)	1.75	2.06	1.75
I42. Facilito que los estudiantes colaboren en la concreción de algunos elementos del sistema de evaluación (productos o actuaciones de aprendizaje a realizar, momentos o fechas de entrega, propuesta de criterios de evaluación, de instrumentos de evaluación, del sistema de calificación).	3	2.67	1.81
I43. Los estudiantes autoevalúan sus productos o actuaciones, bien individual o grupalmente.	2	2.02	1.95
I44. Los estudiantes evalúan los productos o actuaciones de sus compañeros, bien individual o grupalmente.	1	1.98	1.98
I45. Se evalúa y califica de forma dialogada (entre profesorado y alumnado) y consensuada.	1	1.55	1.74

En la Figura 2 se presenta la posición, dispersión y asimetría de las valoraciones realizadas por el profesorado en las cuatro dimensiones. Esta perspectiva global indica que el mayor grado de acuerdo o frecuencia de uso se produce en las dimensiones de profundización (M=4.44) y transparencia (M=4.18) y el menor grado de acuerdo o

frecuencia en las dimensiones de retroalimentación (M=3.30) y participación (M=2.06).

Se ofrecen a continuación (Figura 3) los resultados de cada una de estas cuatro dimensiones de forma independiente.

Figura 2. Diagrama de caja y bigotes correspondiente a las cuatro dimensiones del autoinforme RAPEVA

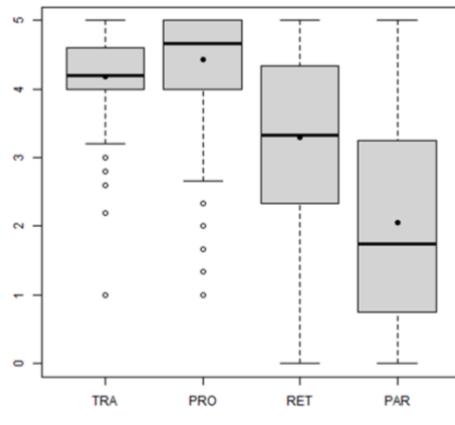
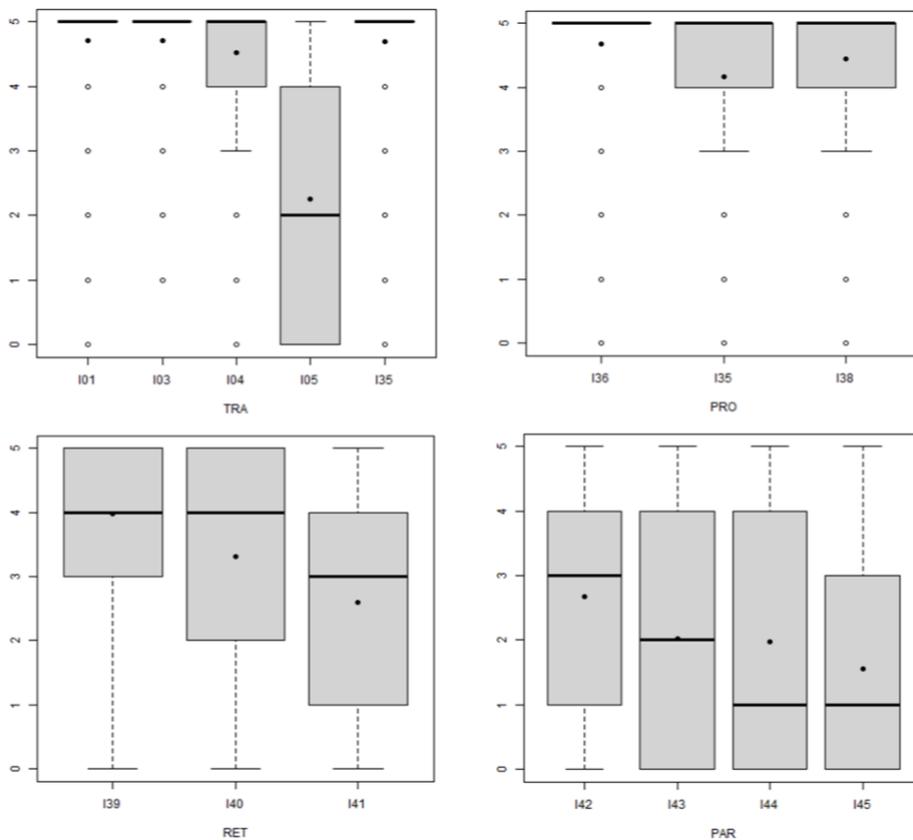


Figura 3. Diagrama de caja y bigotes de los ítems en las cuatro dimensiones



La transparencia en las tareas de evaluación

En la Figura 3 se observan en los ítems de la dimensión de transparencia (TRA) puntuaciones elevadas y concentradas. De hecho, en la Tabla 4 observamos medias superiores a 4.5, a excepción del ítem I05 (M=2.25). El 95.1% del profesorado expresa un alto grado de acuerdo con que “los resultados de aprendizaje, criterios, procedimientos de evaluación y sistema de

calificación de la asignatura son públicos y aparecen en la guía docente o programa” (I01); así mismo el 94.9% considera que “en el sistema de evaluación se informa acerca de los criterios e instrumentos de evaluación” (I03). En este mismo sentido se expresa el 94.7% del profesorado que considera que informa y describe a los estudiantes las tareas de evaluación que tienen que realizar, y ofrecen para ello guías, adendas o transparencias, entre

otros recursos posibles (I35). Un menor grado de acuerdo se presenta con el ítem I04, ya que el 91.9% considera que “En el sistema de evaluación se informa acerca de los criterios e instrumentos de evaluación”. Por último, hay que señalar que existe un bajo grado de acuerdo sobre que el sistema de evaluación informe sobre las diferentes modalidades participativas de evaluación (I05), como la autoevaluación, la evaluación entre iguales o la coevaluación, siendo sólo el 36% del profesorado el que mantiene un alto grado de acuerdo con esta afirmación.

La profundidad de las tareas de evaluación

En referencia a la profundidad de las tareas (PRO) es destacable el elevado grado de acuerdo del profesorado (95,4%) con que las tareas de evaluación que proponen a los estudiantes “implican el uso de conocimientos y contenidos relevantes de la materia” (I36). A continuación, se expresa un 89,4% de acuerdo en relación a que estas tareas propuestas facilitan la aplicación de conocimientos y habilidades a situaciones o casos parecidos a los que encontrarán en su futura vida profesional (I38). Por último, el 79,8% se manifiesta de acuerdo con que las tareas de evaluación se plantean como un reto para los estudiantes (I37).

La retroalimentación en las tareas de evaluación

En relación con la retroalimentación constatamos (Figura 3, RET) cómo se da una mayor dispersión y variabilidad en las opiniones con respecto a las dos dimensiones anteriores. En este caso el 72.1% del profesorado manifiesta que aportan retroalimentación a los estudiantes a través de estrategias como la realización de borradores en los que se aporta una revisión individual, la realización de prácticas o talleres o la comparación con buenas prácticas o productos (I39).

El 53.6% manifiesta que facilita a los estudiantes la participación y la auto retroalimentación, entregándoles previamente instrumentos de evaluación, que les pueden servir para contrastar sus avances con los

estándares explicitados en ellos, y también buenas prácticas o productos con los que comparar los propios (I40).

Por último, sólo el 40.6% considera que los estudiantes reciben retroalimentación de sus propios compañeros a través de la revisión entre iguales de borradores, presentaciones, etc. (I41).

La participación de los estudiantes en la evaluación

En esta dimensión de participación es en la que se presentan puntuaciones medias más bajas en todos los ítems (Tabla 4) y, en todos ellos, por debajo de 3. Como se contempla en la Figura 3 (PAR), se observa una mayor variabilidad en las valoraciones, destacando en primer lugar que sólo el 40.2% manifieste que habitualmente suele favorecer la colaboración de los estudiantes en el diseño de elementos del sistema de evaluación como la concreción de los productos o actuaciones que se considerarán para evaluar, los momentos o fechas de entrega, los propios criterios de evaluación, el diseño de instrumentos o el sistema de calificación (I42). El 30.8% manifiesta que nunca llega a solicitar esta colaboración del estudiantado.

En referencia a la puesta en práctica de modalidades participativas de evaluación, el 48.8% manifiesta no utilizar la autoevaluación (I43) y más de la mitad del profesorado (51.2%) expresa que no utiliza la evaluación entre iguales (I44). En el caso de la coevaluación (I45), el porcentaje de profesorado que no hace uso de esta estrategia llega al 58.7%.

Diferencias perceptivas sobre las tareas de evaluación diseñadas

Una de las cuestiones de investigación de este trabajo hace referencia a las posibles diferencias en función de características del profesorado que pudieran ser útiles para una mejor comprensión de la práctica evaluativa en educación superior.

Un primer análisis multivariante (PERMANOVA I) se realizó considerando

variables dependientes las cuatro dimensiones. Los resultados de este se presentan en la Tabla 5. Se encontraron diferencias significativas en función de la universidad ($F_5, 415=2.29, p=0.016$), el ámbito de conocimiento ($F_2, 415=12.78, p=0.001$) y la satisfacción evaluadora ($F_3, 415=9.5196, p=0.001$). No se hallaron diferencias que estuvieran asociadas a la experiencia docente ($F_2, 415=1.3466, p=0.225$) o al género ($F_2, 415=1.5486, p=0.165$). Para analizar las posibles diferencias significativas en la forma en la que el profesorado diseña las tareas de evaluación a partir de características contextuales como la universidad, el ámbito de conocimiento, la experiencia o la satisfacción con el sistema de evaluación se realizaron análisis comparativos, siguiendo una estrategia de arriba-abajo, utilizando el análisis multivariado permutado de la varianza (PERMANOVA). Se comenzó por analizar las diferencias a partir de las cuatro dimensiones globales para continuar hasta llegar a la comparación de todos los ítems simultáneamente. Los resultados de estos análisis se presentan en la Tabla 5. Por último, para el análisis individual de cada ítem se aplicó la prueba H de Kruskal-Wallis cuyos resultados se exponen en la Tabla 6.

En el análisis post-hoc, mediante la comparación por pares, se constató que las diferencias significativas entre universidades se encontraban entre UNIOVI y UCA ($p=0.015$), así como entre UPV/EHU y UCA ($p=0.002$). No se obtuvieron diferencias significativas en los restantes emparejamientos entre universidades.

En función del ámbito de conocimiento las diferencias se hallaron entre EDU y ECO ($p=0.001$), así como entre ECO y COM ($p=0.003$). Por último, con respecto a la seguridad y satisfacción evaluadora las diferencias se dieron entre el Nivel 3 y el Nivel 1 ($p=0.022$), el Nivel 1 con el Nivel 2 ($p=0.001$), el Nivel 1 con el Nivel 4 ($p=0.001$) y el Nivel 2 con el Nivel 4 ($p=0.004$).

Para un análisis con mayor grado de detalle se procedió a realizar un nuevo análisis (PERMANOVA II) tomando como variables

dependientes cada uno de los ítems del autoinforme que constituían cada una de las dimensiones, lo que permitió analizar las diferencias bajo un patrón completo de múltiples variables dependientes. En la Tabla 5 se presentan los resultados de este análisis multivariante, encontrándose diferencias significativas ($p<.05$) en las cuatro dimensiones en función de la universidad (UNI) y de la satisfacción evaluadora (SSE). Las diferencias por el ámbito de conocimiento sólo se encontraron en las dimensiones RET y PAR. En el caso del género y la experiencia no se encontraron diferencias significativas.

Desde la perspectiva multivariante se realizó un tercer análisis (PERMANOVA III) considerando todos los ítems individualmente como variables dependientes. En este análisis se constataron diferencias significativas ($p<.05$) de los grupos en función de la universidad, el ámbito de conocimiento y la satisfacción evaluadora.

Finalmente, examinando cada uno de los ítems de forma independiente se realizó la prueba H de Kruskal-Wallis (Tabla 6). Se encontraron diferencias significativas ($p<.05$) entre los grupos en función de la universidad en varios ítems, como el I01, I04, I05 e I35 de la dimensión transparencia, en todos los ítems relacionados con la dimensión profundidad, el ítem I39 de la dimensión retroalimentación y el I42 de la dimensión participación en la evaluación. También se observaron diferencias en el ámbito de conocimiento, como el ítem I05 de la dimensión transparencia, el I41 de la dimensión retroalimentación y todos los ítems de la dimensión participación en la evaluación. Además, se encontraron diferencias según el género en el ítem I42 de la dimensión participación en la evaluación, según la experiencia en los ítems I03 e I05 de la dimensión transparencia, y según la satisfacción en la evaluación en todos los ítems analizados, excepto el ítem I01.

La universidad como elemento diferenciador

Las diferencias por universidades se concretan en la dimensión TRA en el ítem I04, referido a ofrecer información sobre las

diferentes modalidades participativas de evaluación. En este ítem las puntuaciones de la UPV/EHU (M=4.27) son menores que las de UNIOVI (M=4.81), UCA (M=4.75) o UDC (M=4.62). Así mismo, en el ítem I05, alusivo a ofrecer información sobre las tareas de evaluación a través de guías, adendas, transparencias, etc., la diferencia significativa se establece entre UPV/EHU (M=1.94) y UCA (M=2.81).

En la dimensión PRO las diferencias entre universidades se presentan en los tres ítems que conforman esta dimensión. Así en el ítem referido a la utilización de tareas de evaluación que implican el uso de conocimientos y contenidos relevantes de las materias (I36), la UPV/EHU presenta las puntuaciones más bajas (M=4.46) frente a las que alcanzan la UNIOVI (M=4.85) y la UCA (M=4.87). Por su parte, en lo que se refiere a la presentación de tareas retadoras para los estudiantes (I37) la diferencia se establece entre la UPV/EHU (M=4.0) y la UCA (M=4.51). Por último, con respecto a la autenticidad de las tareas (I38), es decir las tareas de evaluación que facilitan la transferencia de conocimientos y habilidades a un contexto profesional, la diferencia se presenta entre la UPV/EHU (M=4.18) y la UCA (M=4.59) y UDC (M=4.66).

La diferencia entre universidades en el caso de la dimensión RET radica en el ítem según el cual se proporciona retroalimentación a los estudiantes por parte del profesorado (I39). Las diferencias se establecen entre la URV (M=4.6) frente a la UDC (M=3.64) y la UPV/EHU (M=3.8), y también entre la UV (M=4.3) frente a UDC (M=3.64) y UPV/EHU (M=3.8).

De las diferentes actividades participativas consideradas en la dimensión (PAR), la que establece la diferencia entre las universidades es la referida a la colaboración de los estudiantes en la concreción de algunos de los elementos del sistema de evaluación (I42). En este caso es la URV (M=1.5) la que se diferencia de las restantes universidades UPV/EHU (M=2.62), UDC (M=2.64), UV (M=2.96) y UCA (M=3.21).

Diferencias en función del ámbito de conocimiento

Al abordar las posibles diferencias que el profesorado explicita en sus percepciones sobre la base del ámbito de conocimiento en el que desarrolla su labor docente se confirman diferencias significativas en las dimensiones información, retroalimentación y participación. En el caso de la información (I05), el profesorado de Comunicación (M=2.72) y Educación (M=2.40) se diferencian de los de Economía y Empresa (M=1.99). Esta diferencia también se manifiesta en el I41 entre los ámbitos de Comunicación (M=3.5) y Educación (M=2.89) con respecto a Economía y Empresa (M=2.10).

Por último, respecto a la participación, las diferencias se encuentran en los cuatro ítems de esta dimensión. En el ítem alusivo a la colaboración de los estudiantes en la concreción de los elementos de la evaluación (I42), Comunicación (M=3.33) y Educación (M=3.02) difieren de Economía y Empresa (M=2.11). En la autoevaluación (I43) la diferencia se encuentra entre Comunicación (M=2.67) y Educación (M=3.02) frente a Economía y Empresa (M=1.46). Similar situación nos encontramos con la evaluación entre iguales (I44) donde Comunicación (M=2.67) y Educación (M=2.45) se diferencian de Economía y Empresa (M=1.33) y en el uso de la evaluación dialógica y consensuada (I45), con mayores puntuaciones de Educación (M=1.81) y Comunicación (1.89) frente a las de Economía (M=1.06).

Tabla 5. Resultados de PERMANOVA en función de la universidad, ámbito, género, experiencia y satisfacción con la evaluación

Dimensiones		UNIVERSIDAD			ÁMBITO			GÉNERO			EXPERIENCIA			SATISFACCIÓN		
		F	Sig.	R ²	F	Sig.	R ²	F	Sig.	R ²	F	Sig.	R ²	F	Sig.	R ²
Dimensiones globales	PERMANOVA I	2.2948	.016	.027	12.78	.001	.058	1.5486	.165	.007	1.3466	.225	.006	9.5196	.001	.064
Ítems Dimensión TRA	PERMANOVA II	2.7512	.003	.032	2.2213	.071	.010	1.3037	.258	.006	3.414	.009	.016	4.6197	.001	.032
Ítems Dimensión PRO		2.780	.003	.032	0.6418	.637	.003	0.7603	.480	.003	1.1895	.299	.005	8.1473	.001	.056
Ítems Dimensión RET		1.8399	.044	.022	6.4295	.001	.030	0.7483	.564	.003	2.0355	.087	.009	8.7211	.001	.059
Ítems Dimensión PAR	PERMANOVA III	2.2845	.013	.027	13.792	.001	.062	1.8584	.094	.008	0.5718	.705	.002	3.9533	.001	.028
Ítems RAPEVA		2.2938	.004	.027	8.4893	.001	.039	1.3697	.174	.006	1.5244	.126	.007	5.657	.001	.039

Tabla 6. Resultados de la prueba de Kruskal-Wallis en función de la universidad, ámbito, género, experiencia y satisfacción con la evaluación

	UNIVERSIDAD			ÁMBITO			GÉNERO			EXPERIENCIA			SATISFACCIÓN		
	H	Sig.	η^2	H	Sig.	η^2	H	Sig.	η^2	H	Sig.	η^2	H	Sig.	η^2
Transparencia															
I01	11.064	.050	.027	.407	.816	.001	.513	.774	.001	.968	.968	.000	6.704	.082	.016
I03	4.527	.476	.011	.011	.995	.000	.908	.635	.002	6.758	.034	.016	19.521	.000	.047
I04	26.790	.000	.065	1.294	.524	.003	1.813	.404	.004	5.776	.056	.014	49.785	.000	.120
I05	13.969	.016	.034	6.396	.041	.015	3.420	.181	.008	9.449	.009	.023	9.338	.025	.023
I35	11.770	.038	.028	3.725	.155	.009	3.748	.154	.009	.291	.865	.001	21.021	.000	.051
Profundidad															
I36	26.459	.000	.064	1.593	.451	.004	1.958	.376	.005	1.087	.581	.003	28.092	.000	.068
I37	12.495	.029	.030	.821	.663	.002	2.311	.315	.006	3.982	.137	.010	28.751	.000	.069
I38	16.715	.005	.040	.790	.674	.002	1.246	.536	.003	2.183	.336	.005	33.316	.000	.080
Retroalimentación															
I39	22.159	.000	.053	5.435	.066	.013	.970	.616	.002	1.122	.571	.003	23.642	.000	.057
I40	6.756	.239	.016	4.875	.087	.012	1.520	.468	.004	2.291	.318	.006	44.559	.000	.107
I41	7.138	.211	.017	20.434	.000	.049	.934	.627	.002	5.091	.078	.012	17.730	.000	.043
Participación															
I42	19.585	.001	.047	26.518	.000	.064	7.645	.022	.018	.065	.968	.000	8.344	.039	.020
I43	7.589	.180	.018	24.533	.000	.059	2.606	.272	.006	3.828	.147	.009	9.432	.024	.023
I44	7.003	.220	.017	31.239	.000	.075	3.926	.140	.009	1.337	.512	.003	8.378	.039	.020
I45	9.053	.107	.022	20.538	.000	.049	.280	.869	.001	.876	.645	.002	11.334	.010	.027

La experiencia docente como elemento diferenciador

En el caso de la experiencia tan sólo se han encontrado diferencias significativas entre el profesorado en la dimensión de información en los ítems I03 y I05.

En el ítem I03, es el profesorado con una experiencia de más de 20 años de docencia ($M=4.78$), frente a los que tienen una experiencia menor, de entre 11 y 20 años ($M=4.66$), quienes consideran que en el sistema de evaluación se informa a los estudiantes de lo que tienen que entregar o ejecutar.

Así mismo, se presenta una diferencia entre el profesorado con menos años de experiencia ($M=2.67$) y los de mayor experiencia ($M=2.02$) sobre la información que se ofrece sobre las diferentes modalidades participativas de evaluación como la autoevaluación, la evaluación entre iguales o la coevaluación (I05).

Satisfacción evaluadora como elemento diferenciador

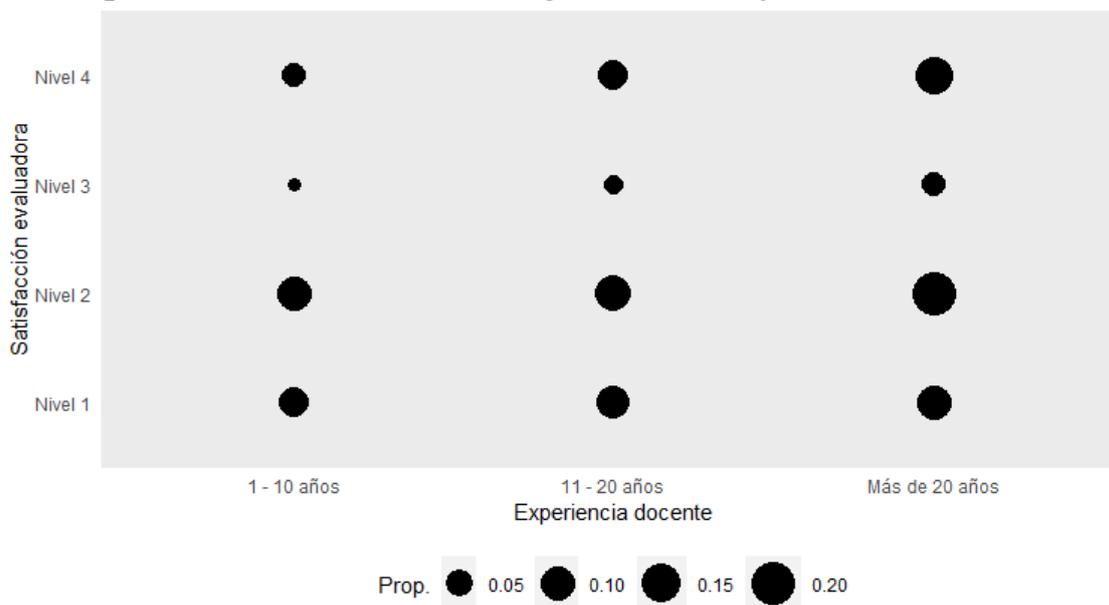
Los niveles de satisfacción evaluadora se han construido sobre la base de las percepciones de los propios profesores en su

respuesta a los indicadores referidos a su seguridad con el sistema de evaluación (I48) y su grado de satisfacción con este sistema (I49). A partir de las puntuaciones alcanzadas en estos dos ítems se consideraron cuatro niveles tomando como punto de corte los cuartiles.

Para comprobar si se da una cierta asociación entre estos niveles de satisfacción y la experiencia docente, se realizó la prueba de Chi-cuadrado y se obtuvo un valor 6.149 ($p=.407$), por lo que se concluye la no existencia de asociación entre estas dos variables. De hecho, el profesorado de cada nivel de experiencia se distribuye en diferentes niveles de satisfacción evaluadora, como se puede observar en el diagrama de burbujas presentado en la Figura 4. Por ejemplo, un 22.1% del profesorado con más de 20 años de experiencia se sitúa en el nivel 1, un 42.6% en el nivel 2, un 7.8% en el nivel 3 y un 27.5% en el nivel 4.

Sobre la base de esta clasificación en función del nivel de satisfacción evaluadora se evidencian diferencias significativas en todos los ítems del autoinforme (Tabla 6) y, en todos los casos, las diferencias más evidentes se establecen entre el nivel 1 y el nivel 4, a favor de este último.

Figura 4. Distribución de niveles de experiencia docente y satisfacción evaluadora



Discusión

Con este estudio se pretendía, en primer lugar, analizar cómo caracteriza el profesorado universitario las tareas de evaluación que diseñan en su práctica evaluativa. En segundo lugar, se pretendía comprobar si existe alguna diferencia en esta caracterización en función de la universidad de pertenencia, el ámbito de conocimiento, el género, la experiencia o la satisfacción evaluadora. Los resultados de este trabajo sugieren implicaciones, tanto desde una perspectiva teórica como práctica y, a su vez, permite vislumbrar líneas futuras de investigación.

Implicaciones teóricas

Una primera aportación de este trabajo es la caracterización de las tareas de evaluación en torno a los cuatro elementos que se han considerado de relevancia. Los resultados obtenidos evidencian que, desde la percepción del profesorado, es práctica común facilitar al estudiantado la información sobre el proceso de evaluación, de forma que el estudiante pueda disponer de toda la información y orientaciones necesarias para conseguir los resultados de aprendizaje deseados. Estos resultados son coherentes con las aportaciones de diferentes autores como Biggs & Tang (2011) y Boud (2022) quienes consideran la necesidad de que los estudiantes estén informados claramente de lo que han de realizar y los estándares mínimos que han de alcanzar, aunque cabe señalar la falta de información aportada a los estudiantes en relación a las modalidades participativas de evaluación, de un modo más destacable por parte del profesorado con más experiencia. Así mismo, se ha destacado cómo el profesorado considera que diseña tareas de evaluación con un alto grado de profundidad, lo que facilita el aprendizaje profundo del estudiantado.

No obstante, se ha podido constatar que el uso de la retroalimentación o, en mayor medida aún, la participación de los estudiantes en el proceso evaluativo, son aspectos ante los cuales el profesorado aún muestra cierta resistencia, a pesar de los evidentes beneficios que ello implica para el estudiantado, como

han evidenciado los trabajos de Panadero et al. (2017), Tai et al. (2018) y las revisiones de Carless & Boud (2018) y Carless & Winstone (2020).

Una segunda implicación hace referencia a la importancia de los aspectos contextuales. En este sentido, los resultados evidencian que la universidad en la que se desarrolla profesionalmente el profesorado, o su ámbito de conocimiento específico, son posibles elementos moduladores de la práctica evaluativa. Pero es la satisfacción con la evaluación la variable con un mayor grado de diferenciación entre el profesorado. Este hecho puede deberse a que el profesorado considera que diseña su actividad evaluadora de forma coherente a su bagaje, conocimiento y experiencia. No obstante, es preciso cuestionar este hecho en tanto, como señalan DeLuca & Johnson (2017), numerosos docentes toman decisiones sobre la evaluación sin la suficiente experiencia o capacitación para ello, y, como afirman Mehrabi & Hosseini (2021), sin un modelo sólido como guía en el que el profesorado pueda apoyarse para diseñar una evaluación orientada a actividades de resolución de problemas de manera colaborativa.

Implicaciones prácticas

Desde una perspectiva práctica, a tenor de los resultados alcanzados, es preciso impulsar y avanzar en una formación del profesorado universitario que persista en la necesaria actitud crítica para favorecer los cambios necesarios, pues tal y como afirman Lo & Leung (2022) para un desarrollo de la evaluación educativa eficaz se hace imprescindible la alfabetización en evaluación del profesorado. El impulso al uso de las tecnologías, a tenor de la pandemia de la COVID-19, hacen más necesaria aún esta actitud crítica ante la modernización y los cambios. La tecnología puede facilitar los procesos de retroalimentación y la participación de los estudiantes en los procesos de evaluación (Gómez-Ruiz et al., 2020; Ibarra-Sáiz & Rodríguez-Gómez, 2017; Rodríguez-Gómez & Ibarra-Sáiz, 2016) lo cual

redundará en la mejora de la reflexión y la autorregulación del estudiantado, pero sólo podrán mejorar su aprendizaje si utilizan de manera adecuada y efectiva la información que se les aporta.

Limitaciones e investigación futura

Metodológicamente este trabajo presenta una serie de limitaciones que, a su vez, constituyen nuevas vías de investigación futura. En primer lugar, el estudio se ha realizado en un contexto específico en universidades públicas españolas y sería necesario ampliar a otras universidades públicas y privadas, así como en otros países lo que permitiría un análisis comparativo. En segundo lugar, es un estudio que se ha realizado a partir de las percepciones del profesorado, por lo que ampliar la comprensión de las prácticas evaluativas mediante el análisis en profundidad a través de un estudio de caso múltiple permitiría analizar en profundidad la naturaleza de las tareas de evaluación diseñadas por el profesorado y cómo éstas son percibidas y experimentadas por el estudiantado.

Así mismo, sería interesante profundizar en la comprensión sobre la satisfacción del profesorado con su práctica evaluativa y en qué medida es crítico con sus actuaciones, pues como señalaba Boud (2020, p. 7) “Un extraño se asombraría al descubrir cuánta práctica evaluativa se lleva a cabo y que no podría ser defendida sobre la base del conocimiento académico”. De esta forma se podría alcanzar una mejor comprensión no sólo de la práctica evaluativa, sino sobre todo de las concepciones que determinan en gran medida esas prácticas, hasta el punto de existir una gran falta de conexión entre las calificaciones otorgadas a los estudiantes y los resultados de aprendizaje esperados para un curso o asignatura (Boud, 2020).

Conclusión

Los resultados de aprendizaje del estudiante pueden ser potenciados si las tareas de evaluación que se diseñen poseen suficiente calidad, es decir, si son transparentes,

potencian el aprendizaje profundo, favorecen la retroalimentación y hacen partícipe al estudiantado en el proceso de evaluación.

En este estudio se ha analizado la percepción que los docentes de distintas universidades tienen sobre las tareas que diseñan. Se ha constatado que, en general, la transparencia, es decir, la información que ha de trasladarse al estudiante para que sea consciente de qué se le va a pedir para alcanzar los resultados de aprendizaje pretendidos, es muy adecuada. De la misma manera, los docentes manifiestan que exigen a sus estudiantes un conocimiento profundo en el que han de establecer relaciones complejas entre conceptos fundamentales implicándoles en tareas contextualizadas en su futura práctica profesional.

Por el contrario, no es tan favorable la situación en lo que se refiere a la retroalimentación y la participación. Los docentes señalan que proporcionan retroalimentación a los estudiantes sobre su progreso en el proceso de enseñanza-aprendizaje. Sin embargo, la participación de los estudiantes en esta retroalimentación sea autoevaluándose o sea evaluando a sus compañeros es todavía muy reducida. En esta misma línea, se sitúa en general la participación de los estudiantes en las tareas de evaluación, dado que no se ofrecen posibilidades de colaboración en la concreción de algunos elementos del sistema de evaluación y tampoco se potencia la evaluación o la calificación de forma dialogada y consensuada entre el profesorado y el alumnado.

Ni el género ni la experiencia docente parecen tener una asociación con la percepción de los docentes sobre su práctica evaluativa. No obstante, sí hay diferencias entre el profesorado de diferentes universidades y ámbitos de conocimiento. En este último caso, el profesorado de los másteres de economía y empresa expresan una menor práctica que los de educación y comunicación en el uso de la retroalimentación o la participación del estudiantado en la evaluación. Por último, el

grado de satisfacción del profesorado con su práctica evaluadora se presenta como determinante. Cuanto mayor es el nivel de satisfacción se tiene una mejor percepción de la calidad de las tareas de evaluación. En cualquier caso, se ha constatado que es necesario impulsar la formación de los docentes para promover el uso de la retroalimentación y, sobre todo, facilitar la participación de los estudiantes en la evaluación, lo que mejoraría la calidad de las tareas de evaluación y redundaría en una mayor consecución de los resultados de aprendizaje deseados.

Financiación y agradecimientos

Este trabajo ha sido posible gracias al Proyecto FLOASS – *Resultados y analíticas de aprendizaje en la educación superior: Un marco de acción desde la educación sostenible* (Ref. RTI2018-093630-B-I00), financiado por el Ministerio de Ciencia, Innovación y Universidades (MCIU) en el Programa Estatal de I+D+i Orientada a los Retos de la Sociedad, la Agencia Estatal de Investigación (AEI) y el Fondo Europeo de Desarrollo Regional (FEDER). Ha contado con el apoyo de la Cátedra UNESCO en *Evaluación, Innovación y Excelencia en Educación* de la Universidad de Cádiz.

Referencias

Anderson, M. J. (2017). Permutational Multivariate Analysis of Variance (PERMANOVA). *Wiley StatsRef: Statistics Reference Online*, 1–15. <https://doi.org/10.1002/9781118445112.stat07841>

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/riee2020.13.1.007>
- 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/riee2022.15.1.008>
- Mehrabi Boshraadi, 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/10.1080/02602938.2019.1600186>
- <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
- 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
- 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^a. Soledad (marisol.ibarra@uca.es)  0000-0003-4513-702X

Profesora Titular de Universidad en el Área de Métodos de Investigación y Diagnóstico en Educación de la Universidad de Cádiz. Directora de la Cátedra UNESCO en Evaluación, Innovación y Excelencia en Educación. Directora del Grupo de Investigación EVALfor-SEJ509 – Evaluación en contextos formativos. Desarrolla su investigación principalmente en el ámbito de la evaluación en educación superior. Ha sido investigadora principal de más de diez proyectos, internacionales y nacionales, cuyos resultados han sido publicados en diversos artículos, capítulos de libros y contribuciones a congresos internacionales. Actualmente es coinvestigadora principal del Proyecto ValeFLOASS – Valorización del entorno e-FLOASS para el seguimiento y evaluación de los resultados de aprendizaje en el sistema educativo (PDC2022-13305-I00) en el que participan nueve universidades españolas.

Lukas-Mujika, José-Francisco (jf.lukas@ehu.eus)  0000-0002-9440-035X

Profesor Titular de Universidad en Evaluación en Educación en la Universidad del País Vasco. Facultad de Educación, Filosofía y Antropología. Departamento de Ciencias de la Educación. Especialista en Evaluación en Educación con especial dedicación a temas relacionados con la evaluación de programas y centros educativos y la construcción de instrumentos de medida. Director Adjunto del Instituto de Ciencias de la Educación de la Universidad del País Vasco (1997-2000). Director del Departamento de Métodos de Investigación y Diagnóstico en Educación de la Universidad del País Vasco (2003-2009). Miembro del Comité Científico del Instituto Vasco de Evaluación e Investigación Educativa no universitaria (ISEI-IVEI). Evaluador en distintas agencias de evaluación y acreditación (ANECA, DEVA, ACSUCYL). Autor de artículos, libros, capítulos de libro y aportaciones en conferencias internacionales.

Ponce-González, Nicolás (nicolas.ponce@uca.es)  0000-0003-1336-6303

Doctor en Ciencias de la Educación por la Universidad de Granada. Profesor Sustituto Interino en el Área de Métodos de Investigación y Diagnóstico en Educación, Universidad de Cádiz. Forma parte del equipo de trabajo de la Cátedra UNESCO en Evaluación, Innovación y Excelencia en Educación. Ha participado en proyectos de investigación a nivel autonómico, nacional e Internacional. Actualmente es miembro del equipo de trabajo del Proyecto ValeFLOASS – Valorización del entorno e-FLOASS para el seguimiento y evaluación de los resultados de aprendizaje en el sistema educativo (PDC2022-133050-I00). Miembro del Grupo de Investigación EVALfor-SEJ509 - Evaluación en contextos formativos. Sus principales líneas de investigación son formación para el desarrollo de competencias, evaluación de programas de intervención en Inteligencia Emocional y Evaluación en Educación Superior. Coautor de varias publicaciones de artículos, capítulos de libro, y comunicaciones a congresos nacionales e internacionales.

Rodríguez-Gómez, Gregorio (gregorio.rodriguez@uca.es)  0000-0001-9337-1270

Catedrático de Universidad de Métodos de Investigación en Educación de la Universidad de Cádiz. Es coordinador del área estratégica "Estudios e investigaciones en evaluación" de la Cátedra UNESCO en Evaluación, Innovación y Excelencia en Educación. Miembro fundador del Grupo de Investigación EVALfor-SEJ509 - Evaluación en contextos formativos. Sus intereses de investigación se centran en la metodología de investigación y en la evaluación en educación superior. Actualmente es coinvestigador principal del Proyecto ValeFLOASS – Valorización del entorno e-FLOASS para el seguimiento y evaluación de los resultados de aprendizaje en el sistema educativo (PDC2022-133050-I00) en el que participan nueve universidades españolas. Autor de artículos, capítulos de libro y aportaciones en conferencias internacionales. Ha sido presidente de la Asociación Interuniversitaria de Investigación Pedagógica (AIDIPE) y de la Red Estatal de Docencia Universitaria 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/).