

EDUCACIÓN, CREATIVIDAD E INTELIGENCIA ARTIFICIAL: NUEVOS HORIZONTES PARA EL APRENDIZAJE. ACTAS DEL VIII CONGRESO INTERNACIONAL SOBRE APRENDIZAJE, INNOVACIÓN Y COOPERACIÓN, CINAIC 2025

María Luisa Sein-Echaluce Lacleta, Ángel Fidalgo Blanco y Francisco José García Peñalvo (coords.)

1ª Edición. Zaragoza, 2025

Edita: Servicio de Publicaciones. Universidad de Zaragoza.



Servicio de
Publicaciones
Universidad Zaragoza

EBOOK ISBN 978-84-10169-60-9

DOI 10.26754/uz.978-84-10169-60-9



Esta obra se encuentra bajo una licencia Creative Commons Reconocimiento – NoComercial (ccBY-NC). Ver descripción de esta licencia en <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Referencia a esta obra:

Sein-Echaluce Lacleta, M.L., Fidalgo Blanco, A. & García-Peñalvo, F.J. (coords.) (2025). *Educación, Creatividad e Inteligencia Artificial: nuevos horizontes para el Aprendizaje. Actas del VIII Congreso Internacional sobre Aprendizaje, Innovación y Cooperación. CINAIC 2025 (11-13 de Junio de 2025, Madrid, España)*. Zaragoza. Servicio de Publicaciones Universidad de Zaragoza. DOI 10.26754/uz.978-84-10169-60-9

Sesgos en la evaluación continua

Biases in continuous assessment

Antonio Álvarez Fernández-Balbuena, Ricardo Bernárdez Villaboa, Daniel Vazquez Moliní, Francisco Díez Garrote
antonioa@ucm.es, rbvoptom@fis.ucm.es, dvazquez@ucm.es, fdiez@ucm.es

Dpto. Óptica. F. de Óptica y Optometría
UCM
Madrid, España

Dpto. Óptica. F. de CC Físicas
UCM
Madrid, España

Resumen- Se analizan los sesgos en la evaluación continua en la asignatura "Instrumentos Ópticos y Optométricos" en la Universidad Complutense de Madrid. Se estudian tres cursos académicos (2019-2022) con diferentes profesores para determinar si la evaluación continua es equitativa. Los resultados muestran diferencias significativas entre los grupos A, B y C, a pesar de compartir exámenes finales. La evaluación continua representa el 15% de la nota final, y se concluye que este método puede perjudicar a ciertos grupos de estudiantes. Se sugiere revisar los métodos de evaluación para garantizar la justicia en la calificación. El estudio destaca la importancia de la coordinación entre profesores y la necesidad de criterios de evaluación claros y consensuados.

Palabras clave: Evaluación Continua, Sesgos, Instrumentos Ópticos, Equidad, Criterios de Evaluación.

Abstract- Biases in continuous assessment are analyzed in the subject "Optical and Optometric Instruments" at the Complutense University of Madrid. Three academic years (2019-2022) with different professors are studied to determine if continuous assessment is fair. The results show significant differences between groups A, B, and C, despite sharing final exams. Continuous assessment represents 15% of the final grade, and it is concluded that this method may disadvantage certain groups of students. It is suggested to review the assessment methods to ensure fairness in grading. The study highlights the importance of coordination among professors and the need for clear and agreed-upon evaluation criteria.

Keywords: Continuous Assessment, Biases, Optical Instruments, Fairness, Evaluation Criteria

1. INTRODUCCIÓN

En este complejo mundo de la docencia, existen multitud de condicionantes que influyen en el resultado final obtenido por los alumnos.

Normalmente se tiende a focalizar desde casi todos los ámbitos, instituciones, responsables de calidad e incluso medios de comunicación, a la evaluación como un sistema de medición, pero en muchas ocasiones sistemas de evaluación fiables y rigurosos se acompañan de aprendizajes superficiales y por lo tanto de corta duración. (Gibbs y Simpson, 2005)

Ya en la década de los 70 se investigaba en prestigiosas universidades cómo era el aprendizaje de los estudiantes, encontrando que la actividad de estos, como la asistencia a

clase, horas de trabajo, etc., estaba plenamente condicionada a las demandas del sistema de evaluación, estudios posteriores cualitativos destacan la importancia de entender como responden los estudiantes a las innovaciones en la evaluación. (Snyder, 1971; Miller y Parlett, 1974; Sambell y McDowell, 1998).

Diversos estudios afirman que las calificaciones en los trabajos tienden a ser mejores que en los exámenes, hay estudios que las sitúan en torno al 12% más altas y otros que ponen de manifiesto que la mayoría de los alumnos prefieren que al menos la mitad de la nota provenga de actividades y trabajos durante el curso, los alumnos también consideran que la evaluación basada en estos trabajos y actividades es más justa que los exámenes. Sin embargo, el aprendizaje basado en estos trabajos y actividades no supone que éste sea más superficial o profundo por parte de los alumnos. (Gibbs y Simpson, 2005)

La calidad de la implicación de los estudiantes en las actividades de aprendizaje resulta crucial. La cantidad de trabajos, calificaciones, no tiene por qué generar el aprendizaje necesario, por el contrario, otras estrategias como ofrecer una guía e incrementar la orientación a los estudiantes en sus trabajos y un feedback adecuado y en tiempo si ayudan. Calificaciones sin feedback pueden ser dañinas y afectar a la autoeficacia (capacidad y esfuerzo) por lo que este debería estar centrado en el contenido, dando opciones al estudiante para actuar, al no estar centrado en él mismo sino en su actuación. (Gibbs y Simpson, 2005)

La calificación es percibida en ocasiones por el estudiante como un indicador de su habilidad personal con respecto a otros. Teniendo en cuenta que la relación entre el esfuerzo y la calificación no es siempre directa, la percepción de los estudiantes de su esfuerzo, depende más de la motivación que del número de horas, horas que son muchas veces improductivas de aprendizaje superficial, asociado al tipo de evaluación. Algunos estudios concluyen que mantener la motivación es la cuestión más importante. (Kember et al. 1996: Gibbs y Simpson, 2005)

En relación a la calificación los estudiantes necesitan entender por qué han recibido la calificación o el comentario y por qué no han tenido uno más alto o más bajo. Criterios explícitos y comprendidos por los estudiantes que queden patentes en las calificaciones ayudan estos a saber que estándar se considera el adecuado. (Gibbs y Simpson, 2005)

Los sistemas de evaluación continua ofrecen al profesorado la oportunidad de hacer un seguimiento preciso del proceso de aprendizaje de los estudiantes y obtener evidencias de los resultados alcanzados. Las actividades de evaluación son ocasiones privilegiadas para proporcionar apoyo, seguimiento y tutorización. Los criterios de evaluación deben ser consensuados en cada área de conocimiento o departamento y más concretamente entre los profesores implicados en la docencia, pero la presencia de condicionantes asociados a los profesores y a los propios alumnos hacen de estos criterios no perfectamente objetivos. (Gibbs y Simpson, 2005)

En la universidad ha predominado desde siempre una cultura profesional individualista, la cual poco a poco se va superando con grupos de profesores que trabajan en proyectos educativos conjuntos. Algunos estudios señalan que casi el 53% de los profesores se coordina con los compañeros en aspectos de contenido, metodología y evaluación de los programas. (Feixas, 2004)

Actualmente la organización de los departamentos sigue siendo diversa, encontrando tipos, como el académico tradicional, el gerencial y el departamento académico como equipo, que sería el más deseado para mejorar la coordinación. (Ramsden, 1998)

Entre las dificultades de coordinación a superar, se encuentra la falta de oportunidades para el profesorado joven, para los que, generalmente los compañeros más veteranos son referencias de los que se puede aprender, pero muchas veces estos veteranos desconfían de lo que puedan aportar los jóvenes. Cada departamento debe convertirse en un espacio de aprendizaje donde compartir conocimiento, preocupaciones y experiencia, sin estar exento de un apoyo formativo externo. Feixas, 2004)

El desarrollo de la carrera docente de cada profesor universitario se modifica como respuesta a experiencias profesionales, personales, (etapas vitales, relaciones familiares, etc.) y organizativas (expectativas sociales, regulaciones de profesión. etc.) (Burke, 1987)

En cada profesor por lo tanto podríamos hablar de: Variables personales; Las de tipo docente, como la planificación, estrategias, recursos de evaluación, concepciones de la enseñanza y del aprendizaje; Variables de tipo relacional o su relación con estudiantes y compañeros; De tipo institucional, como la cultura docente y departamental; Y de la variable investigación en la orientación docente. Todas ellas a su vez se interrelacionan con la edad, años de experiencia docente o formación inicial entre otras. La motivación del profesor sin duda alguna está influenciada por todas estas variables. (Biggs, 2000)

La influencia educativa de cada profesor es esencial tanto a hora del aprendizaje como de la evaluación, puede ser útil para cuantificarla fijarse en los diferentes tipos de ayudas que el docente ofrece a los estudiantes y cómo estos las valoran: guías didácticas; material de ayuda; plataforma Moodle; comunicación presencial con el profesor; etc. (Feixas, 2004)

2. CONTEXTO Y DESCRIPCIÓN

La evaluación de asignaturas en el ámbito universitario para alcanzar los objetivos definidos en el Verifica de la Aneca requiere cumplir estrictamente con los rangos de evaluación de distintos apartados (Universidad Complutense de Madrid,

2016). Concretamente dentro de la Facultad de Óptica y Optometría se imparte el Grado en Óptica y Optometría y diversos Másteres. El objetivo de la titulación es formar a especialistas en óptica y optometría. Dentro del Grado se pueden distinguir dos perfiles fundamentales, el perfil optométrico y el perfil de óptica técnica. Las competencias se han desarrollado para estudiantes que han superado los estudios de bachillerato y/o similares y los han orientado hacia la rama donde se enmarca esta titulación. Estas competencias se basan en las que aparecen en la Orden Ministerial por la que se establecen los requisitos para la verificación de los títulos que habiliten para el ejercicio de la profesión de Óptico Optometrista. Todas estas competencias se plantean desde la base de la igualdad entre hombres y mujeres, la igualdad de oportunidades y accesibilidad a personas con discapacidad y los valores democráticos y los propios de una cultura de la paz. Estas competencias garantizan el Marco Español de Calificaciones para la Educación Superior MECES.

No se han encontrado, en la bibliografía, trabajos que aborden este tipo de estudios.

Este estudio va a tener como objetivo la evaluación continua en la asignatura Instrumentos Ópticos y Optométricos que en la UCM es obligatoria con una carga de 6 créditos ECTS y tiene lugar en el segundo curso y primer semestre formando parte del módulo de Óptica. En el marco actual la docencia comprende presencialmente:

- 52.5 h de clases teoría (30 h), seminarios (15 h), trabajos tutelados (7.5 h)
- 10 h de clases de laboratorio (4 prácticas de 2.5 h)

A. Evaluación

La evaluación permite medir el desempeño académico de los estudiantes y el grado de adquisición de las competencias necesarias. Se proponen diferentes actuaciones para hacer la evaluación, examen final, prácticas y evaluación continua. Una media ponderada de estas evaluaciones dará la nota final del alumno que deberá ser mayor o igual a cinco para optar al aprobado.

En el Verifica tenemos los siguientes rangos con los que decidir, Tabla 1:

Tabla 1: Rangos permitidos para evaluación en Verifica

SISTEMA DE EVALUACIÓN	PONDERACIÓN MÍNIMA	PONDERACIÓN MÁXIMA
Evaluación de conocimientos teóricos.	50	75
Evaluación de conocimientos prácticos y/o problemas	10	20
Participación en seminarios y trabajos personales	10	20

Considerando estos rangos la propuesta para la calificación final del alumno es:

- Examen final de la asignatura: 75 %
- Evaluación de prácticas: 10 %
- Evaluación continua: 15 %

B. Evaluación continua

En este trabajo se va a analizar el sesgo que puede tener la evaluación continua dependiendo del profesor, se van a analizar tres cursos académicos 2019-2020, 2020-2021, 2021-2022 en

los que la evaluación continua se ha realizado por el mismo profesor en cada uno de los grupos y años. Con esta información el objetivo de este trabajo es determinar si en este caso concreto hay que buscar alternativas a la evaluación continua o pensar en modificar el porcentaje de peso en la calificación final. Cada profesor ha realizado un mínimo de 4 pruebas de evaluación y un máximo de 30 según su propio criterio, los tres profesores conocen los ejercicios de los demás para ajustar el nivel de conocimiento e intentar adecuarlos en el mismo nivel.

Debido a que la evaluación continua se realiza en horario de clase y cada grupo A, B y C tienen horarios distintos que no se solapan cada profesor hace sus evaluaciones en su clase y no son las mismas para cada grupo. Los grupos A y B son de mañana y el C de tarde.

3. RESULTADOS

El grupo consta de 621 sujetos repartidos en tres cursos académicos. Cada curso académico está dividido a su vez en tres grupos, A, B y C cada uno de ellos impartido por un mismo profesor durante los tres cursos.

Empleando el software de estadística SPSS versión 29.0.2.0 (20) se analiza primero la diferencia de las medianas para la evaluación continua (Figura 1) comprobando que existe diferencia significativa entre grupos ($P < ,05$) para la evaluación continua dependiendo del profesor, se ha analizado para tres cursos académicos 2019-2020, 2020-2021, 2021-2022 donde esta evaluación continua se ha realizado por el mismo profesor en cada uno de los grupos y años.

N total	621
Mediana	4,933
Estadístico de prueba	151,964
Grado de libertad	2
Sig. asintótica (prueba bilateral)	<,001

Figura 1. Resumen de prueba de mediana de muestras independientes

En la Figura 2 se observa la mediana de evaluación continua de los grupos A, B y C juntando los tres cursos académicos.

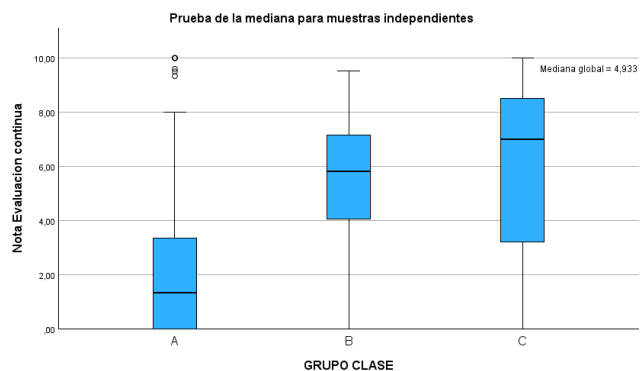


Figura 2. Prueba de mediana para muestras independientes. 3 cursos juntos.

Siguiendo el estudio de los tres cursos a la vez, en la Figura 3, se observa la comparativa por grupos de clase, donde los

valores de significación mediante la corrección de Bonferroni, indican diferencias significativas A-B y A-C. Es decir, el grupo A tiene una diferencia respecto al resto de grupos.

Sample 1-Sample 2	Estadístico de prueba	Sig.	Sig. ajust. ^a
A-B	130,108	<,001	,000
A-C	105,861	<,001	,000
B-C	8,038	,005	,014

Figura 3. Comparaciones por parejas de grupo clase.

Anualmente se puede comparar si entre grupos A, B y C la diferencia en la evaluación continua es significativa como se puede observar en la Figura 4 para el Grupo A.

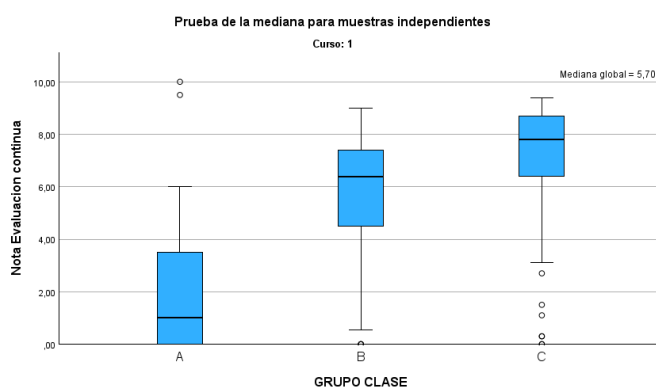


Figura 4. Notas EC curso 1

Tabla 2. Estadísticas EC. Datos de significación entre los grupos en curso 1

Sample 1-Sample 2	Estadístico de prueba	Sig.	Sig. ajust. ^a
A-B	49,624	<,001	,000
A-C	99,866	<,001	,000
B-C	20,765	<,001	,000

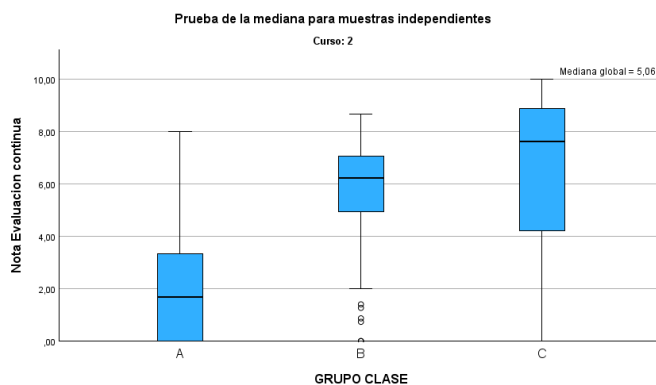
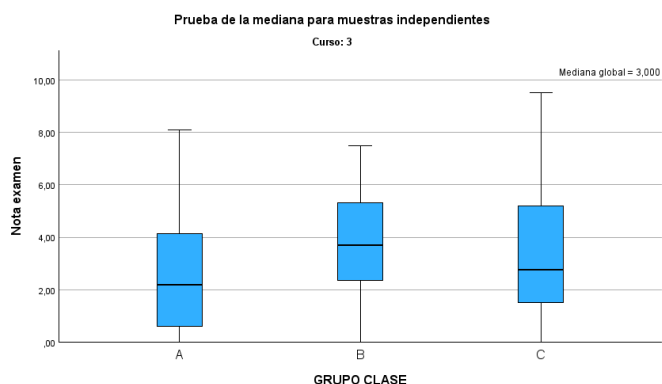


Figura 5. Notas EC Datos de significación entre los grupos en curso 2

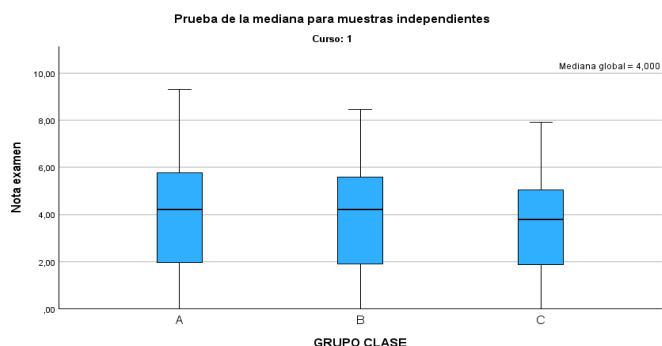
Tabla 3. Estadísticas EC curso 2

Sample 1-Sample 2	Estadístico de prueba	Sig.	Sig. ajust. ^a
A-B	43,638	<,001	,000
A-C	36,546	<,001	,000
B-C	6,833	,009	,027

**Figura 6.** Notas EC Datos de significación entre los grupos en curso 3**Tabla 4.** Estadísticas EC curso 3

Sample 1-Sample 2	Estadístico de prueba	Sig.	Sig. ajust. ^a
A-C	3,805	,051	,153
A-B	8,616	,003	,010
C-B	2,427	,119	,358

Si observamos las notas de examen del curso 1 no se obtiene significación estadística, Figura 7. Con valores por grupos de significación ajustada mediante la corrección Bonferroni en función de la nota de examen en 0.872 y 1.000. El nivel de significación es de 0.05.

**Figura 7.** Notas examen curso 1

4. CONCLUSIONES

La evaluación continua tiene un peso del 15% sobre la nota final. En este trabajo se demuestra, para una asignatura específica y durante tres años consecutivos, que la evaluación continua no es equivalente en cuanto a nota obtenida por el alumno en cada grupo con distintos profesores. A pesar de compartir exámenes para equilibrar el nivel de dificultad. Se

puede descartar que la EC sea menor en algún grupo de alumnos debido a un nivel más bajo puesto que en el examen final los grupos A, B y C obtienen notas significativamente iguales.

Por ello, es necesario estudiar métodos con los que ser justos en la evaluación continua pues en este caso una ponderación del 15% sobre la nota perjudica a un grupo de alumnos. En la tabla 5 se proponen algunas opciones para evitar sesgos.

Tabla 5. Propuestas de mejoras para evitar sesgos

Mejoras para evitar sesgos en formación continua por grupos		
Rúbrica común entre grupos de la misma asignatura	Revisión cruzada entre docentes	Pruebas compartidas

AGRADECIMIENTOS

Subvención proyecto del Ministerio de Ciencia, Innovación y Universidades proyectos de Generación de Conocimiento 2022 PID2022-138061OB-I00 “Tecnologías ópticas avanzadas aplicadas al análisis y restauración de tapices históricos”

REFERENCIAS

- Biggs, J. (2000). Teaching for quality learning at university. Buckingham: The Society for Research in Higher Education and Open University Press
- Burke, P.J. (1987). Teacher Development. Induction, Renewal and Redirection. London: The Falmer Press
- Feixas, M. (2004) La influencia de factores personales, institucionales y contextuales en la trayectoria y el desarrollo docente de los profesores universitarios. *Educación* 33, 31-54.
- Gibbs, G., & Simpson, C (2005) Conditions under which assessment supports students' Learning. *Learning and Teaching in Higher Education* 1. 3-31.
- Kember, D., Ng, S., Tse, H., Wong, E.T.T. & Pomfret, M. (1996) An examination of the interrelationships between workload, study time, learning approaches and academic outcomes, *Studies in Higher Education*, vol. 21, no. 3, pp. 347-358.
- Miller, C.M.I. & Parlett, M. (1974) Up to the Mark: a study of the examination game, Guildford: Society for Research into Higher Education
- Ramdsen, P. (1998). Learning to lead in higher education. Londres: Routledge
- Sambell, K. & McDowell, L. (1998) The construction of the hidden curriculum: messages and meanings in the assessment of student learning, *Assessment and Evaluation in Higher Education*, vol. 23, no. 4, pp. 391-402
- Snyder, B.R. (1971) The Hidden Curriculum, Cambridge, MA: MIT Press.
- Universidad Complutense de Madrid. (2016). *Memoria de Verificación ANECA (Grado)*. [https://optica.ucm.es/data/cont/docs/13-2016-03-08-1%20Memoria%20de%20Verificaci%C3%B3n%20ANECA%20\(Grado\).pdf](https://optica.ucm.es/data/cont/docs/13-2016-03-08-1%20Memoria%20de%20Verificaci%C3%B3n%20ANECA%20(Grado).pdf)