

# Evaluación formativa entre pares. Experiencias en todos los niveles universitarios

## ***Autores***

*Diez-Fernández, Ángeles\*, Alonso-Cortés Fradejas, M. Dolores, Barba-Martín, Raúl, Domínguez-Fernández, Raquel, Fernández-Fernández, Roberto, García-González, Marta Eva, González-Alonso, M. Inmaculada, Gutiérrez-García, Carlos, Herrero-Molleda, Alba, Hortigüela-Alcalá, David, Pérez-Pueyo, Ángel*

*\*Departamento de Filología Moderna*

## ***Nombre del Grupo de Innovación***

*IFAHE. Innovation through Formative Assessment in Higher Education*

## **RESUMEN**

El objetivo de este trabajo es describir tres tipos de experiencia de evaluación entre pares diseñados con el fin de mejorar no solo el aprendizaje de las materias, sino también las competencias profesionales tanto de evaluados como de evaluadores. Otro de los fines fue estimular la reflexión de los docentes implicados para que, tras un debate sobre las ventajas e inconvenientes del proceso, generasen protocolos y repertorios de buenas prácticas transferibles entre asignaturas y campos de conocimiento. El primer tipo de experiencia, llevada a cabo entre estudiantes en grupos pequeños o medianos en once asignaturas de grado y máster, consistió en la evaluación de trabajos en grupo escritos y/o de sus exposiciones orales. La segunda, realizada entre pares de alumnos en gran grupo, se centró en pruebas “clásicas” escritas utilizadas como parte de la evaluación formativa de una asignatura. Por último, en la tercera experiencia profesores expertos evaluaron la actuación docente de sus pares noveles. Tras concluir las fases de preparación, evaluación y recogida de resultados de las mismas, y a falta del análisis final de los datos de las encuestas de satisfacción y el rendimiento académico, podemos decir que los resultados parecen prometedores: se ha producido un aumento de la motivación y la familiarización con instrumentos de evaluación como las escalas de valoración. Los participantes parecen haber mejorado también sus competencias comunicativas, su capacidad reflexiva y “habilidades blandas” (*soft skills*) como la capacidad crítica y la colaboración, además de avanzar en su autonomía personal y profesional. Como conclusión, podemos decir que este es un paso adelante en la búsqueda de experiencias de evaluación eficaces desde el punto de vista formativo, sostenibles para el sistema universitario, el docente y la asignatura, y transferibles entre asignaturas de todos los niveles educativos.

**Líneas de actuación:** Aprendizaje Colaborativo; Innovación en Evaluación

## **Introducción**

En contextos universitarios ya debería haber sido superada la idea tradicional del feedback como la información que el docente proporciona a sus estudiantes con el fin de que estos mejoren las calificaciones de sus producciones, entre otros motivos por el carácter poco estable de sus resultados (no siempre mejoran el aprendizaje ni los resultados académicos) y su difícil gestión tanto a nivel práctico (consumen mucho tiempo) como afectivo (las críticas no se aceptan de forma positiva).

Desde enfoques docentes más innovadores centrados en el alumno, el feedback se describe como un proceso (no un producto) mediante el que los aprendices dan sentido a la información recibida de

fuentes diversas y la utilizan para mejorar sus trabajos y sus estrategias de aprendizaje. En este marco es necesario que el evaluado sepa qué es el feedback y cómo gestionarlo de forma adecuada. Además, debe valorar el papel del docente y el suyo propio dentro del proceso (Carless & Boud, 2018).

Una de las técnicas más eficaces para conseguir un feedback efectivo es integrar dentro de la evaluación formativa procesos de evaluación entre pares, un acuerdo en el que los individuos consideran la cantidad, nivel, valor, calidad o éxito de los productos o resultados de aprendizaje alcanzados por individuos del mismo estatus o similar (Topping, 1998). Cuando un individuo valora los trabajos de sus compañeros, puede evaluar los suyos de forma más eficaz, al poder comparar ambos. Es más, a menudo proporcionar comentarios a los pares es más beneficioso que recibirlos, al exigir niveles de implicación cognitiva más altos, como son la aplicación de criterios, el diagnóstico de problemas o la sugerencia de soluciones (Nicol et al., 2014). Por último, dado que la interacción entre los participantes es una condición necesaria, la participación en estos procesos incentiva el desarrollo de sus destrezas comunicativas y de colaboración, además de mejorar la capacidad de discriminar entre productos, los hábitos reflexivos y la autonomía (Ocampo et al., 2022).

Sin embargo, y dado que la evaluación formativa entre pares debe entenderse como una experiencia colaborativa, con el fin de maximizar sus beneficios los diseñadores de los procesos deben tener en cuenta cuestiones afectivas que, muchas veces, generan el rechazo inicial de los futuros evaluadores a participar en la experiencia. La primera de ellas es su inseguridad sobre el conocimiento de la materia y las destrezas evaluadoras propias y de sus pares. La segunda tiene que ver con la gestión de sus emociones y sus relaciones en el aula, a veces, conflictivas (Hamodi, 2016).

Para la superación de estos retos, Carless y Boud (2022) insisten en la necesidad de desarrollar lo que ellos denominan *student feedback literacy* (alfabetización del feedback del estudiante), que incluye cuatro elementos fundamentales: (1) formar a evaluados y evaluadores para que aprecien el feedback y entiendan la necesidad de implicarse activamente en el proceso, (2) entrenarlos para que sean capaces de emitir juicios sobre la calidad del trabajo propio y del de otros, (3) apoyarlos en la gestión de las emociones derivadas de emitir y/o recibir juicios positivos o negativos y (4) animarlos a actuar y a utilizar de forma constructiva los comentarios en sus producciones posteriores. Para todo ello es necesario que los participantes reciban formación e instrumentos de evaluación adecuados (Mulder et al., 2014), así como espacios y tiempos para desarrollar su aprendizaje sobre la evaluación.

Este trabajo presenta diversas experiencias de evaluación entre pares, diseñadas por los componentes del grupo IFAHE en sus asignaturas siguiendo las cuatro premisas mencionadas. y puestas en práctica dentro de sistemas de evaluación formativos más amplios.

## Experiencia innovadora

Se describen tres tipos de experiencias de evaluación entre pares, de características diferentes: (1) Evaluación por pares de estudiantes de trabajos en grupo escritos y/o de sus exposiciones orales, (2) Evaluación por pares de estudiantes de pruebas “clásicas” escritas utilizadas como parte de la evaluación formativa, y (3) Evaluación por pares de la actuación docente de profesores noveles por parte de profesores expertos.

### *Objetivos*

El objetivo principal fue implementar procesos de evaluación entre pares que mejorasen no solo el aprendizaje de las materias, sino también, y, sobre todo, las competencias profesionales de los participantes. Se intentaba, además, estimular la reflexión de los docentes que habían diseñado y desarrollado los diferentes procedimientos. El fin último era que estos pusiesen en común y discutiesen las ventajas e inconvenientes encontrados en el proceso y generasen un protocolo y/o repertorio de buenas prácticas transferible entre asignaturas y campos de conocimiento que, además, les permitiese desarrollar su trabajo de forma más eficaz y motivadora.

### *Descripción de la experiencia*

Los tres tipos de experiencias, que se implementaron durante tres fases idénticas (preparación, evaluación y recogida y análisis de resultados), se describen a continuación.

#### 1. Primera experiencia: la evaluación por pares de exposiciones orales y escritas

Este tipo de experiencia se llevó a cabo en once asignaturas de siete grados de diferentes en dos universidades (León y Burgos) y en cuatro del Máster Universitario en Formación del Profesorado en ESO y Bachillerato, Formación Profesional y Enseñanzas de Idiomas (MUFPEs). Estas asignaturas fueron las siguientes: *Aprendizaje de la Comunicación Escrita y Aprendizaje de las Ciencias Sociales* (4º de Educación Infantil), *Aprendizaje y Enseñanza de las Ciencias Experimentales II: Biología y Geología y Prácticum II: Mención Inglés* (3º y 4º de Educación Primaria, respectivamente), *Diseño y Análisis de la Enseñanza de la Actividad Física* (4º de Ciencias de la Actividad Física y el Deporte), *Educación Física y su Didáctica* (2º de Educación Primaria, Universidad de Burgos), *Ingeniería Eléctrica* (4º de Ingeniería Aeroespacial), y *Fundamentos de los Contenidos de la Educación Física y Prácticum* de módulos de Geografía e Historia, Inglés y Tecnologías del MUFPEs.

Aunque con peculiaridades dentro de las diferentes asignaturas, estos procesos de evaluación formativa entre pares siguieron tres fases comunes:

1. *Fase de preparación.* Los docentes presentaron en el aula los objetivos y características del trabajo oral y/o escrito y del tipo de evaluación programada con el fin de descubrir al alumnado el potencial de la evaluación por pares para la mejora de

los resultados académicos y del nivel competencial tanto de evaluadores como de evaluados, y mejorar su motivación e implicación en el proceso. Además, los implicados empezaron a familiarizarse con los instrumentos de evaluación. En todas las asignaturas se utilizó una escala de valoración con los criterios de evaluación y niveles de logro correspondientes. Las diferencias básicas entre ellas en esta fase fueron dos. En unos casos la autoría de la escala fue responsabilidad solo del docente, mientras que en otros fueron los propios estudiantes, orientados por el profesor, quienes la elaboraron. La segunda especificidad afecta a la profundidad del trabajo previo con la escala. Mientras que, en ocasiones, la presentación de la misma fue más breve (lectura y solución de dudas puntuales), en otras se realizaron actividades como el andamiaje y el modelado, “ensayando” el uso de la rúbrica y su protocolo de aplicación, por ejemplo, evaluando trabajos de estudiantes de cursos anteriores, o sesiones previas de formación sobre cómo proporcionar un feedback adecuado y efectivo.

2. *Fase de evaluación.* Los estudiantes utilizaron la escala para tomar notas de los puntos fuertes y débiles de los trabajos de sus pares y proporcionarles feedback. Dado que estos procedimientos se integraron dentro de procesos de evaluación formativa, sus contribuciones permitieron a los compañeros mejorar la calidad del mismo producto evaluado o de los consecutivos en una misma asignatura. Una de las diferencias entre las experiencias realizadas se encuentra en el tipo de feedback ofrecido una vez concluidas las presentaciones orales o escritas: en general, este fue oral y grupal, cuando toda la clase ofreció sugerencias de mejora a quien acababa de exponer, o individual y escrito, presentando los resultados mediante la rúbrica tras un período de reflexión. Otra de las variaciones se relaciona con el número de veces que se repitió la experiencia para “preparar” el producto final, que osciló entre tres y siete veces. La última particularidad fue que, en algunos casos, la coevaluación se utilizó sola, actuando el docente como mero mediador, mientras que en otros de acompañó de procesos de autoevaluación (cuando se pidió a los propios evaluados que utilizaran la escala para valorar sus productos) y/o heteroevaluación (cuando también el profesor evaluó el producto).

3. *Fase de recogida y análisis de los resultados.* Se utilizaron métodos distintos para recoger la información sobre el nivel de satisfacción y percepción del aprendizaje de los implicados. Mientras en la mayoría de las asignaturas se realizó mediante encuestas virtuales (a través de Moodle o Google Forms), en otras se hizo mediante diálogos semi-estructurados docente-estudiantes en el aula.

## 2. Segunda experiencia: Evaluación por pares de pruebas escritas “clásicas” en gran grupo

Esta experiencia fue una prueba piloto realizada en la asignatura *Lengua extranjera aplicada a la enseñanza I: Inglés*, de 1º de Educación Primaria. Implementar procesos de evaluación formativa en este contexto presenta varias dificultades, como el elevado número de estudiantes (130), su nivel (recién llegados a la universidad), la heterogeneidad del dominio previo del contenido o las características de la asignatura (se evalúan tanto las destrezas receptivas como las productivas). La experiencia que se presenta aquí utilizó la evaluación por pares como complemento a la autoevaluación y a la heteroevaluación, y sus fases fueron las siguientes:

1. *Fase de preparación.* La docente presentó en el aula los objetivos de la evaluación formativa y el procedimiento para llevarla a cabo: se realizarían entre diez y doce pruebas escritas cortas (cinco preguntas) de comprensión lectora y oral durante el semestre en las que los alumnos deberían justificar sus respuestas escribiendo unas dos o tres líneas. El texto de las respuestas se utilizaría para evaluar formativamente la destreza de expresión escrita. Para motivarles se les explicó que este tipo de trabajo podría mejorar su aprendizaje y sus calificaciones finales.

2. *Fase de evaluación.* La labor de la profesora fue realizar una detección rápida de los errores en los textos redactados, subrayándolos e indicando, mediante un código compartido de símbolos, si la pregunta estaba bien o mal respondida, total o parcialmente. Después, las devolvió a los alumnos, que iniciaron un proceso de reflexión y autoevaluación intentando encontrar por sí mismos las respuestas adecuadas. Si no tenían éxito se procedía a desarrollar la evaluación por pares, en la que eran los compañeros quienes les daban el feedback necesario. La profesora intervino durante este proceso solo en momentos puntuales.

3. *Fase de recogida y análisis de resultados.* Los alumnos respondieron a una encuesta de satisfacción y percepción del aprendizaje (Google Forms).

## 3. Evaluación por pares de profesores expertos de la actuación docente de colegas noveles

Esta experiencia comenzó en 2021 con el objetivo de guiar y apoyar en la adquisición de competencias docentes en la universidad a una profesora novel (beneficiaria de una ayuda FPU) en la Facultad de Ciencias de la Actividad Física y del Deporte, donde imparte docencia en la asignatura de *Biomecánica de la Actividad Física y del Deporte*. Se está iniciando, además, con otra doctoranda del área de Didáctica de la Lengua y la Literatura del Departamento de Filología Hispánica y Clásica.

Las fases de la experiencia, que se prevé que se siga desarrollando de forma cíclica (en concreto las fases 2 y 3) hasta el final de su ayuda FPU, son las siguientes:

1. *Fase de preparación.* La profesora novel asistió a las clases de la asignatura que iba a dar, impartidas por el responsable de la misma, experto en Biomecánica.

Además, recibió formación sobre la utilización del diario de aula y del “Registro de observación de la actuación docente”, un instrumento de evaluación a nivel de sesión adaptado y testado en experiencias similares previas (González-Alonso et al., 2019), a partir del registro de observación de la Universidad de Sheffield utilizado para la acreditación docente de su profesorado.

2. *Fase de evaluación.* Cuando ella fue la encargada de la docencia, periódicamente asistieron a sus clases evaluadores, miembros de IFAHE, que disponían de los datos previos, como el plan de clase, con los objetivos, contenidos y metodología, entre otros elementos, compartidos por la evaluada en su registro. En una reunión posterior los evaluadores proporcionaron a su evaluada feedback sobre la secuencia docente presenciada, señalando las fortalezas y debilidades detectadas y proponiendo posibles mejoras que la profesora novel intenta integrar en sus próximas clases.

3. *Fase de recogida y análisis de los resultados.* Tras finalizar cada ciclo, se analiza el conjunto de producciones elaboradas (diario de la profesora y registro de observación), complementándose con entrevistas a los diversos agentes implicados.

En los tres tipos de experiencias, en los que se han generado, además de los documentos con los diseños y programaciones, diversas escalas de valoración y registros y diarios de aula (estos dos en la última descrita), nos encontramos en la última fase, la del análisis de los resultados recogidos y su puesta en común.

#### *Plan de seguimiento e indicadores para evaluar los resultados*

Durante el curso 2022-2023, los componentes del grupo IFAHE, además de analizar los resultados de las encuestas de satisfacción de los alumnos participantes, pondrán en común toda la información extraída de las diferentes experiencias. El objetivo es definir unas pautas generales y/o un repertorio de buenas prácticas, para el diseño, el desarrollo y la valoración de procesos de evaluación por pares que sean eficaces desde el punto de vista formativo, sostenibles en el tiempo y transferibles a los distintos niveles y campos universitarios. Intentaremos encontrar respuesta a preguntas como: ¿cuántas fases deben realizarse y qué técnicas son las adecuadas en cada una de ellas?, ¿cómo se puede “entrenar” a los alumnos para que evalúen los productos de sus compañeros?, ¿podemos establecer un esquema general para las escalas de valoración? o ¿qué instrumentos son más adecuados para la recogida de datos? Además de los datos cualitativos, cuando sea posible utilizaremos también datos cuantitativos, como la evolución de las calificaciones en las distintas fases de la evaluación formativa o aquellos derivados de entrevistas y técnicas de investigación similares.

## Resultados alcanzados

A falta del análisis de los resultados finales podemos avanzar algunos resultados, fruto de la observación de los procesos en el aula y de las conversaciones con los implicados. Una consecuencia inmediata ha sido la mejora de la asistencia y motivación de los estudiantes en las sesiones de exposición de los trabajos (por la necesidad de evaluarse unos a otros). Además, se han familiarizado con el uso de rúbricas y escalas de valoración, lo que les resultará útil a corto plazo en sus estudios, por ejemplo, en la realización de sus TFG o TFM. Hemos comprobado, también, que han mejorado no solo sus competencias comunicativas escritas y orales (profesores veteranos incluidos), al tener que proporcionar un feedback comprensible y útil a sus compañeros, sino su capacidad reflexiva y sus “habilidades blandas” (*soft skills*): se han acostumbrado a realizar y recibir críticas fundamentadas y a colaborar con otros de forma satisfactoria. Por último, podemos decir que este tipo de evaluación les ha permitido avanzar en su autonomía personal y profesional, lo que es especialmente relevante para los futuros maestros de Educación Primaria y profesores de Educación Secundaria, que necesitan mejorar sus habilidades docentes y están trabajando en la construcción de su identidad profesional.

El punto débil de este proceso se encuentra en los diferentes niveles de implicación de los alumnos. La evaluación por pares depende del compromiso de cada participante. Cuando no ha sido óptimo tanto evaluadores como evaluados se han quejado de que los comentarios emitidos o recibidos son insuficientes (las contribuciones son pocas o demasiado breves), poco adecuados (por la dureza del tono empleado, por ejemplo) o poco útiles (por la falta de reflexión y profundidad, lo que imposibilita la mejora futura). Entendemos que una formación inicial más intensa o la utilización de incentivos como la calificación de la participación podrían mejorar el proceso.

## Conclusiones y valoración de la experiencia

En la búsqueda de formas de llevar las metodologías activas colaborativas a la evaluación, los componentes del grupo IFAHE estamos desarrollando experiencias evaluadoras entre pares, tanto de alumnos como profesores. En ellas, intentamos que todos valoren y agradezcan el trabajo que supone a sus pares proporcionar feedback para la mejora de su aprendizaje de un modo no competitivo. Asimismo, formamos a los implicados para que puedan emitir juicios fundamentados y útiles con afán constructivo y sin miedo al rechazo, apoyados en herramientas para la evaluación fiables, y organizamos secuencias en las que puedan utilizar lo aprendido como evaluados y evaluadores para la mejora de su rendimiento. Además, para los docentes que han diseñado los procesos, una verdadera comunidad de aprendizaje de profesores, este es un paso más en la búsqueda de experiencias de evaluación eficaces desde el punto de vista formativo, sostenibles para el sistema universitario, el docente y la asignatura, y transferibles entre materias de todos los niveles educativos.

## Agradecimientos

Este trabajo ha sido financiado por la Universidad de León dentro del Plan de Apoyo a los Grupos de Innovación Docente. GID 009 IFAHE.

## Referencias bibliográficas

- Carless, D., & Boud, C. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment and Evaluation in Higher Education*, 43(8), 1315-1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Gonzalez-Alonso, M. I., Domínguez-Fernández, R., Díez-Fernández, Á., Pérez-Pueyo, Á., & Gutierrez-García, C. (2019). Autoevaluación y coevaluación de la práctica docente del profesorado universitario. *Infancia, Educación y Aprendizaje*, 5(2), 353-358. <https://doi.org/https://doi.org/10.22370/ieya.2019.5.2.1724>
- Mulder, R. A., Pearce, J. M., & Baik, C. (2014). Peer review in higher education: Student perceptions before and after participation. *Active Learning in Higher Education*, 15(2), 157-171. <https://doi.org/10.1177/1469787414527391>
- Nicol, D., Thomson, A., & Breslin, C. (2014) Rethinking feedback practices in higher education: a peer review perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102-122. <https://doi.org/10.1080/02602938.2013.795518>
- Hamodi, C. (coord.). (2016). *Formar mediante la evaluación en la universidad. Propuestas prácticas útiles para docentes*. UVA. <http://uvadoc.uva.es/handle/10324/42970>
- Ocampo, J. C., Panadero, E., & Díez, F. (2022). Are men and women really different? The effects of gender and training on peer scoring and perceptions of peer assessment. *Assessment & Evaluation in Higher Education*, 1-17. <https://doi.org/10.1080/02602938.2022.2130167>
- Topping, K. (1998). Peer Assessment between Students in Colleges and Universities. *Review of Educational Research*, 68(3), 249-276. <https://doi.org/10.3102/00346543068003249>