

# Análisis de validez en la evaluación entre iguales: un estudio en educación superior

*Analysis of validity of peer assessment: a study in Higher Education*

**Pascual-Gómez, Isabel; Lorenzo-Llamas, Elena María & Monge-López, Carlos**

Universidad de Alcalá de Henares

---

## Resumen

No son muy numerosas las investigaciones que se han realizado en España sobre la participación de los estudiantes en los procesos de *evaluación entre iguales* en el marco de la Educación Superior. El objetivo de este trabajo es conocer con mayor precisión un proceso de evaluación participativa en el marco de un proyecto de aprendizaje cooperativo. Se recogieron datos de una muestra de 119 estudiantes matriculados en una asignatura del Grado de Magisterio en la Universidad de Alcalá. El procedimiento de recogida de datos fue online. Se diseñaron actividades de evaluación entre iguales y autoevaluación grupal. Los datos se compararon con las calificaciones emitidas por el profesor. Los resultados del análisis de contraste de medias arrojan una diferencia significativa de 1-2 puntos entre ambas puntuaciones a favor del alumnado. A pesar de esta sobrevaloración, el análisis de contenido del instrumento de autoevaluación concluye la utilidad de esta investigación: este tipo de procesos mejora *capacidad formativa de los procesos de evaluación* tanto para el discente como para el docente.

**Fecha de recepción**  
21 de Julio de 2014

**Fecha de aprobación**  
18 de Enero de 2015

**Fecha de publicación**  
3 de Abril de 2015

## Palabras clave:

Aprendizaje cooperativo; autoevaluación; carpeta de aprendizaje; Educación Superior; evaluación entre iguales; evaluación para el aprendizaje; heteroevaluación; perspectiva émica; rúbrica

---

## Abstract

In Spain only a limited number of studies have been done regarding students evaluating their peers in higher education. The objective of this article is to analyse the way in which students that have participated in a cooperative learning project evaluated their peers. Data were collected online from a sample of 119 bachelor students that were registered in a teacher training course at the University of Alcalá. Activities were designed for both peer evaluation and self-evaluation in group work. These student evaluations were compared with ratings given by the teacher. A contrast analysis showed a statistically significant difference of 1-2 points between the mean scores, in favour of the students. Despite these overestimated values, an analysis of the content from the self-evaluation tool showed its usefulness for improving the formative assessment process for both students and teachers.

**Reception Date**  
2014 July 21

**Approval Date**  
2015 January 18

**Publication Date:**  
2015 April 3

## Keywords:

Cooperative learning; self-assessment; portfolio assessment; Higher Education; peer-assessment; assessment for learning; hetero-assessment; emic perspective; rubric

---

Una de las tareas fundamentales del profesorado es planificar y ejercer evaluaciones, bien sea de sus alumnos, de sus programas o de su propia práctica docente. Una adecuada evaluación hace posible la

mejora continua de la calidad de los procesos de enseñanza-aprendizaje. Por este motivo, en el contexto de la Educación Superior se intenta garantizar una formación apropiada en cuanto a competencias para evaluar se refiere en los

---

*Autor de contacto / Corresponding author*

**Pascual-Gómez, Isabel** ([isabel.pascualg@uah.es](mailto:isabel.pascualg@uah.es)). Departamento de Ciencias de la Educación de la Universidad de Alcalá de Henares. C/ Madrid, 1 19001 Guadalajara (España)

planes de estudios de los estudiantes de Magisterio.

En el contexto internacional numerosas investigaciones se han ocupado de contrastar la validez y fiabilidad de este tipo de evaluaciones. Entre otras, Montero-Fleta (2006) menciona: "las experiencias de evaluación del trabajo en grupo de Freeman (1995), el análisis de los criterios de corrección en la evaluación compartida investigada por Orsmond, Merry y Reiling (1996) y las diferencias de la puntuación otorgada por el profesor y asignada por el alumno de Kwan y Leung (1996) y MacAlpine (1999)" (p. 32).

También pueden señalarse los trabajos de revisión meta-analítica realizados por Bretones Román (2008). Gran parte de estos estudios han corroborado la fiabilidad de este tipo de evaluaciones y el alto grado de acuerdo existente entre las puntuaciones de los alumnos con las del profesor encontrándose correlaciones positivas. Pero también se han encontrado evaluaciones discrepantes (Pond, Ul-Haq & Wade, 1995). Estos autores afirman que en ocasiones se producen sobrevaloraciones por relaciones de amistad, falta de diferenciación dentro del grupo, personas que dominan en el grupo y que incrementan su puntuación y personas que fallan en la contribución al grupo y que se benefician del trabajo de los demás.

Convencidos de que la evaluación debe ser un proceso participativo y cooperativo y considerando que los entornos virtuales de evaluación facilitan este tipo de interacción, este estudio pretende explorar las fortalezas y debilidades de una experiencia de evaluación en un marco de trabajo cooperativo implementada en una plataforma de aprendizaje virtual. Se contrastarán las valoraciones realizadas desde diferentes tipos de evaluación a través de un análisis de validez convergente. Por ello, es necesario aclarar previamente la terminología relativa al propio proceso de evaluación desarrollado en esta investigación para hacerlo más comprensible al lector. Los conceptos teóricos básicos

utilizados en este artículo hacen referencia, principalmente, a dos términos que tienen que ver con modalidades de evaluación según quién evalúa. Para Rodríguez Gómez e Ibarra Sáiz (2011) se pueden diferenciar dos tipos: *heteroevaluación* y *autoevaluación*. En una publicación posterior (Rodríguez Gómez, Ibarra Sáiz & García Jiménez, 2013) llegan a decir: "el prefijo `hetero´ o `auto´, referido a la evaluación, viene determinado por el tipo de relación que mantiene quien evalúa con el objeto que se evalúa (aprendizaje, enseñanza, etc.)" (p. 200).

En este trabajo, la *heteroevaluación*, es la evaluación en la que participan profesor-alumno siendo el profesor el agente que evalúa al alumno y su aprendizaje, con una clara diferenciación de roles. En sentido contrario, la evaluación del alumno sobre el papel del profesor, según señala López Pastor (2009) sería un proceso de evaluación compartida orientada a la mejora de la formación y la docencia del profesor pero no sería *heteroevaluación*.

La *autoevaluación* se define como la evaluación que una persona el estudiante realiza sobre sí misma, su propio aprendizaje o su propio proceso. En la *autoevaluación*, como indica Falchikov (1986), el estudiante juzga su propio rendimiento y/o producto atendiendo a sus propios criterios de evaluación. Falchikov y Boud (1989) comentan las bondades de este tipo de evaluación por proporcionar a los estudiantes la participación en la toma de decisiones sobre su propio aprendizaje. Roberts (2006), valora esta modalidad por considerarla un proceso en el que los estudiantes "tengan reflexión crítica, registros del proceso, y tal vez sugieran grados para su propio aprendizaje" (p. 3). De forma más completa, Rodríguez Gómez, Ibarra Sáiz y García Jiménez (2013) definen la *autoevaluación* como el "proceso mediante el cual los estudiantes realizan un análisis y valoración de sus actuaciones y/o sus producciones". (p. 202). Desde la perspectiva educativa suele referirse a la *autoevaluación* de los principales agentes implicados en el

proceso de enseñanza-aprendizaje, alumno y profesor. No obstante, mayoritariamente se refiere al primero de los agentes: el alumno. Asimismo, Panadero y Alonso Tapia (2013) matizan este concepto entendiendo que conlleva un carácter enriquecedor para el aprendizaje del estudiante, como “un proceso que el alumno realiza para autorregular su aprendizaje [...] y como estrategia de instrucción” (p. 173). En el presente estudio se ha aplicado con carácter grupal más que individual. Obteniendo sus resultados como producto de un consenso llevado a cabo por negociación de las individualidades, es decir, negociación entre los integrantes de un grupo de aprendizaje cooperativo: el alumnado.

La participación del alumnado en los procesos de evaluación es valorada desde estudios previos que ponderan tanto sus beneficios como su necesidad, según se ha visto en primer lugar en este apartado. Desde esta perspectiva positiva, se ha incluido en este trabajo el estudio de la *evaluación entre iguales* como máximo exponente de dicha participación del alumnado. Siguiendo a Falchikov (1995), “la evaluación entre iguales es el proceso mediante el cual los grupos de individuos califican a sus compañeros” (p.175), y así se ha desarrollado este tipo de evaluación en esta investigación. De esta forma, tal y como precisa Topping (2009), los estudiantes han considerado y especificado el nivel, valor, cualidad de un producto (prácticas de la asignatura) y rendimiento de otros estudiantes de igual estatus (compañeros de su grupo de aprendizaje cooperativo).

Asimismo, todos los tipos de evaluación señalados hasta el momento, han de orientarse en este trabajo hacia la consecución de un mismo objetivo académico: el aprendizaje del alumno. De aquí surge un marco teórico-práctico valorado en esta investigación: la *evaluación para el aprendizaje*. Según López Pastor (2009) dicha evaluación hace referencia a una orientación clara a la mejora del aprendizaje del alumno en vez de estarlo al mero control de éste y su calificación. Con una perspectiva más amplia, autores como Boud y

Falchikov (2006) hacen referencia a procesos de desarrollo de capacidades de autorregulación en el propio aprendizaje y de aprendizaje a lo largo de la vida. Conviene señalar que la definición de Boud y Falchikov (1989) es matizada por Brown y Glasner (2003) distinguiendo la promoción de la autorregulación y el control sobre el propio aprendizaje con el apelativo de evaluación formadora. Esta perspectiva se escapa a los objetivos y las limitaciones del presente estudio, por lo cual solo ha sido trabajada la evaluación para el aprendizaje tal y como la definen López Pastor (2009) y Alonso-Tapia y Hernansaiz Garrido (2013). Esto es, como una herramienta pedagógica que influye en el modo en el que los alumnos tratan de aprender (Alonso-Tapia & Hernansaiz Garrido, 2013) y puede favorecer el desarrollo de competencias de aprendizaje permanente (López Pastor, 2009).

Por ello, es necesario señalar las *características de la evaluación para el aprendizaje* que han servido de guía a este estudio. Partimos de los errores que según Sanmartí (2007) cometen los docentes y causan las dificultades en el aprendizaje del alumnado: formas de percibir nuestro entorno; formas de razonar; formas de comunicar; formas de sentir (creencias, valores, emociones); y, las ideas transmitidas. Así pues, las características de una verdadera evaluación para el aprendizaje se sintetizan en: planteamiento de actividades que permitan identificar el origen de las dificultades encontradas en el proceso de aprendizaje; contemplación de los procesos implicados en la comprensión de ideas; y, proporcionar a los alumnos las ayudas adecuadas para que puedan superar sus dificultades de aprendizaje. De acuerdo con Alonso-Tapia y Hernansaiz Garrido (2013), “[...] la evaluación tiene como función primordial identificar el origen de las dificultades de los alumnos en relación con la competencia de que se trate con el fin de poder ofrecer las ayudas específicas necesarias para que estos progresen” (p. 8).

## Método

El objetivo principal de este estudio es obtener información sobre el desarrollo de un proceso de aprendizaje cooperativo y los procesos de evaluación que tienen lugar, en el marco de los estudios universitarios para la formación de maestros. Así, planteamos unos objetivos específicos:

- Conocer la validez convergente, a través del cálculo del grado de acuerdo de la evaluación realizada por los agentes implicados en el proceso de enseñanza-aprendizaje: docente (heteroevaluación) y discentes (autoevaluación y evaluación entre iguales).
- Dar voz a los datos, contrastando una perspectiva de valoración émica (Cochran-Smith & Lytle, 2002).

## Procedimiento

El plan de investigación al que se ajusta este estudio corresponde a un análisis descriptivo por auto-observación con utilización de técnicas mixtas de recogida y análisis de datos. Se implementaron técnicas de recogida de datos que se ajustan a: auto-registros de conductas externas (modelo de Rúbrica creado para la recogida de datos en la evaluación entre iguales y escala de valoración para la heteroevaluación); auto-registros de conductas internas (cuestionario de preguntas abiertas creado para la autoevaluación grupal).

El procedimiento se desarrolló según los criterios establecidos para llevar a la práctica la evaluación entre iguales por Ibarra Saiz, Rodríguez Gómez y Gómez Ruíz (2012). Se implementaron las siguientes etapas:

- a. Una fase de preparación en la cual se discuten los criterios de evaluación referentes a la participación en el grupo así como los instrumentos de evaluación.
- b. Creación de los grupos de aprendizaje cooperativo.

- c. Realización de dos propuestas de trabajo a realizar en los grupos para desarrollar el contenido de la asignatura.
- d. Asignación de roles y reparto de tareas dentro de los grupos de aprendizaje cooperativo.
- e. Desarrollo de las prácticas sobre el contenido de la asignatura y evaluación de éstas. La evaluación del proceso es realizada por los alumnos aportando información sobre cómo realizan la tarea y la implicación de los participantes (autoevaluación grupal). La evaluación del producto es llevada a cabo tanto por el profesor (escala de valoración) como por el alumnado. Se utiliza el Aula Virtual como recurso para garantizar el anonimato en cada una de estas evaluaciones. Controlando así la variable sinceridad en las respuestas dadas.

La sistematización del proceso de auto-observación de esta investigación se llevó a cabo mediante la toma repetida de datos en cada una de las prácticas realizadas para trabajar los contenidos de la asignatura y en el seguimiento del proceso de evaluación (heteroevaluación, autoevaluación y evaluación entre iguales) con los instrumentos creados.

## Participantes

Participaron en este estudio 119 estudiantes de primer curso del Grado en Magisterio de Educación Infantil de la asignatura “Diagnóstico Pedagógico y Técnicas de Observación en el Aula de Infantil” y una profesora. La composición de los grupos generales (grupo-clase de alumnos A, B y C) y subgrupos (29 subgrupos en total, son grupos de aprendizaje cooperativo) aparece reflejada en la Tabla 1.

Tabla 1.- Distribución del alumnado por grupos generales y subgrupos

	N	Número subgrupos	Media de alumnado por subgrupo
Grupo A	41	10	4.1
Grupo B	42	9	5.0
Grupo C	36	10	4.6

### **Instrumentos**

Se utilizaron como instrumentos de evaluación e instrumentos de recogida de datos para la investigación: una Rúbrica para realizar los procedimientos de evaluación entre iguales; un cuestionario para efectuar un procedimiento de autoevaluación grupal negociada diseñado ad hoc; y una escala de valoración para la heteroevaluación, creada ad hoc. Los 3 instrumentos estuvieron disponibles desde el inicio del curso en el Aula Virtual – campus online que utiliza la plataforma *Blackboard Learning System*–. Los instrumentos fueron discutidos en el aula en sesiones presenciales. De esta forma los alumnos desde el comienzo del curso tuvieron una información precisa y detallada sobre su evaluación y una comunicación más fluida durante el proceso.

### **Instrumentos para la evaluación entre iguales**

La Rúbrica valoraba al grupo de trabajo en una serie de categorías que influyen en el trabajo cooperativo. Ésta se elaboró siguiendo el *Esquema de Evaluación del Grupo* propuesto por Heathfield (2003). Los indicadores que la componen son:

- Asistencia regular a las reuniones del grupo
- Aportación de ideas respecto a un tema.
- Material de investigación.
- Análisis y preparación del tema.
- Apoyo y motivación de los miembros del grupo.
- Contribución práctica al producto final.

La Rúbrica combinaba en una escala cuantitativa y/o cualitativa los criterios que permitían valorar el nivel del aprendizaje del alumnado y los niveles progresivos de ejecución en cada tarea. Se propusieron 3 categorías de valoración (0-5-10) para describir los niveles de ejecución.

### **Instrumentos para la autoevaluación**

Para realizar la autoevaluación grupal se consideró más adecuado elaborar un cuestionario ad hoc de preguntas abiertas que permitiese a los alumnos reflexionar y justificar, desde una perspectiva émica (Cochran-Smith & Lytle, 2002), su proceso de trabajo y aprendizaje dentro del grupo. El cuestionario incluía un primer bloque de preguntas abiertas en el cual debían identificar los roles que cada miembro había tenido en el grupo; y, un segundo bloque en el que debían identificar las fortalezas y las dificultades encontradas en el desarrollo de las actividades de evaluación, así como los acuerdos y decisiones tomadas en grupo. Por último, después de realizar estas reflexiones, los integrantes del grupo de trabajo cooperativo debían llegar a un consenso sobre la calificación final (en base 10) del conjunto las actividades realizadas por el grupo en la carpeta de aprendizaje –apreciación del nivel de aprendizaje cooperativo alcanzado–.

### **Instrumentos para la heteroevaluación**

Con el fin de poder realizar una triangulación de las fuentes de información, se elaboró una escala de valoración ad hoc en la cual se establecían los indicadores que el profesor tendría en cuenta a la hora de evaluar los trabajos realizados por el grupo. Esta escala se puso a disposición de los alumnos en el Aula Virtual al principio del semestre para que tuviesen conocimiento previo de ellos y

podiesen tenerlos en cuenta antes de realizar las actividades grupales. Los indicadores fueron:

- Número de actividades entregadas.
- Organización, desarrollo y claridad de las prácticas.
- Reflexiones sobre la aportación de la práctica y conexión con aspectos teóricos.
- Consultas bibliográficas o tipo de argumentaciones realizadas.
- Aspectos formales: Presentación.

Éstos indicadores llevaban asociados cuatro niveles de ejecución (Muy bueno/Bueno/Apto/Repetir), escala tipo Likert de grado 4.

### Análisis de datos

Como técnicas de selección y análisis de datos se implementaron las siguientes:

*Análisis cuantitativo de datos* a través del análisis estadístico, mediante la utilización del software SPSS Statistics v.19.0. Se realizó un estudio de muestras relacionadas. Este análisis incluye descriptivos, un análisis correlacional, un contraste de medias relacionadas (T test) y el tamaño del efecto ( $r$  y  $d$  de Cohen).

*Análisis cualitativo de datos* a través del análisis de contenido, mediante la utilización del software NVivo v.10.0. Se efectuó un análisis ajustado a las categorías emergentes relacionadas con el marco teórico del estudio referido a la evaluación para el aprendizaje (Alonso-Tapia y Hernansaiz Garrido, 2013; Carless, Joughin, Liu, & col., 2006; Gallego,

Sánchez, y Cubero, 2011; Otero, Yuste, y Alzás, 2011; Rodríguez Gómez e Ibarra Saiz, 2011; Sanmartí, 2007). Para evitar la arbitrariedad subjetiva de la categorización dicho proceso fue desarrollado por dos investigadores de manera independiente, en un primer momento, y con una puesta en común final. De esta forma, se establecieron dos grandes categorías –familias de categorías– que son: *aspectos positivos a destacar* y *dificultades encontradas*.

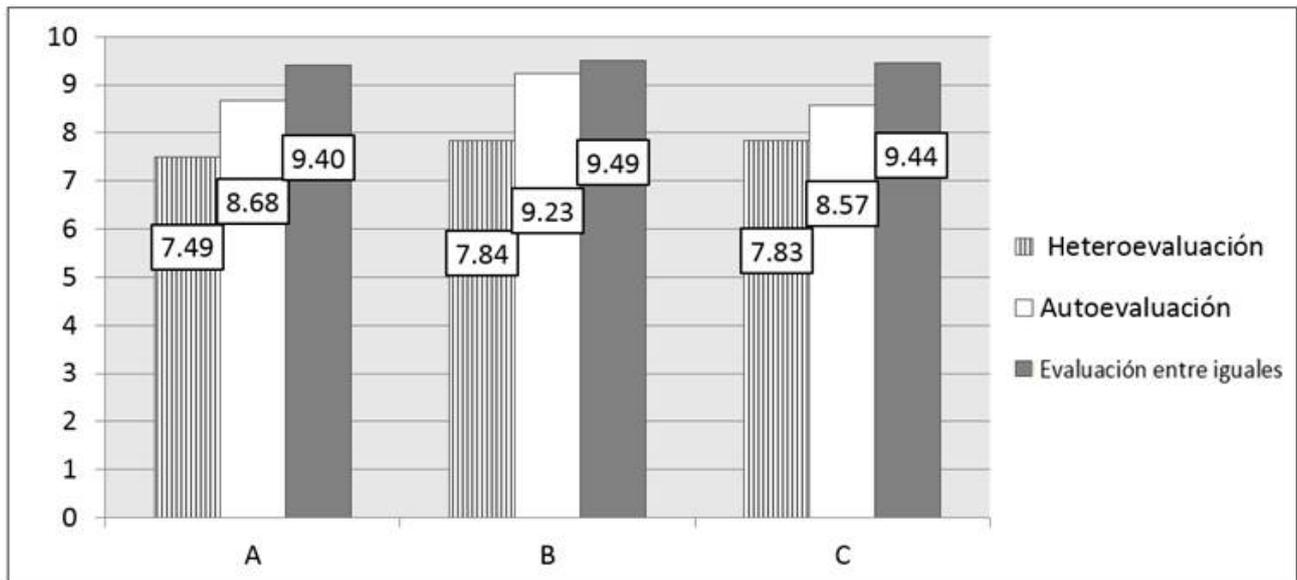
Se controló la reactividad mediante los auto-registros que realizaron los participantes –profesor y alumnado– sin que los investigadores interfirieran.

## Resultados

### Sobre el grado de acuerdo entre las distintas modalidades de evaluación: validez convergente

Como puede observarse en el Gráfico 1 las puntuaciones medias otorgadas por los alumnos a sus propios compañeros fueron altas, (entre 9.40 y 9.49 puntos sobre un total de 10) no existiendo diferencias estadísticamente significativas entre los grupos ( $F=0.59$ ,  $p=.85$ ). Las puntuaciones resultantes de la autoevaluación negociada también son altas (entre 8.57 y 9.23 puntos). Y en este caso, con diferencias estadísticamente significativas y a favor del grupo B, que puntuó más alto que el resto de los grupos ( $F=4.39$   $p=.00$ ). Las calificaciones otorgadas por el profesor son inferiores (entre 1.5 y 2 puntos por debajo) y con una mayor dispersión. Las diferencias entre grupos, en este último caso, no son estadísticamente significativas ( $F=.99$   $p=.37$ ).

Gráfico. 1. Medias de las calificaciones en las tres modalidades de evaluación según los grupos de la muestra



Respecto a la valoración de los indicadores de trabajo cooperativo el aspecto más valorado por el alumnado es *el Apoyo y motivación a los procesos del grupo* (9,1 puntos) siendo los *Materiales aportados en las reuniones y sesiones* el aspecto menos valorado (8.4 puntos).

La Tabla 2 muestra el grado de asociación existente entre las puntuaciones otorgadas por

los alumnos a sus compañeros en una tarea (evaluación entre iguales), la puntuación otorgada por el profesor (heteroevaluación) y la puntuación negociada por el propio grupo (autoevaluación). En todas estas combinaciones las correlaciones encontradas son débiles y no significativas en gran parte.

Tabla 2.- Correlaciones de muestras relacionadas

	N	r	p
Evaluación entre iguales y Heteroevaluación	119	,201	,028
Evaluación entre iguales y Autoevaluación	119	,152	,099
Heteroevaluación y Autoevaluación	119	,022	,812

La Tabla 3 muestra el contraste de medias entre estas mismas puntuaciones. En este caso

se comprueban diferencias estadísticamente significativas en los tres casos.

Tabla 3.- Contraste de medias (prueba t para medidas relacionadas) en las distintas evaluaciones y cálculo del estadístico de Cohen

	Media	Desviación estándar	Error típico	t	gl	Sig.	d
Evaluación entre iguales -Heteroevaluación	1,72	1,34	,12	14,006	118	,00	1.64
Evaluación entre iguales -Autoevaluación	,60	1,19	,10	5,527	118	,00	0.97
Heteroevaluación- Autoevaluación	-1,12	1,61	,14	-7,567	118	,00	0.66

Según Morales Vallejo (2008) una diferencia estadísticamente significativa no es necesariamente una diferencia grande y tampoco es necesariamente una diferencia importante. Por este motivo decidimos realizar el cálculo del tamaño del efecto. Este procedimiento tiene como objetivo la cuantificación de la relevancia del efecto obtenido. Es decir, se trata de establecer si efectos estadísticamente significativos son relevantes en el campo de aplicación de la investigación. Para calcularlo, en el caso que nos ocupa, se calculó el estadístico *d* de Cohen que posteriormente fue transformado en coeficiente de correlación. Se obtuvo un coeficiente de correlación de tipo medio para la diferencia entre la evaluación entre iguales y la heteroevaluación ( $r = .64$ ), más bajo para la diferencia entre la evaluación entre iguales y la autoevaluación ( $r = 0.43$ ) y más moderado en el caso de la diferencia entre autoevaluación y la heteroevaluación ( $r = .31$ ). A la luz de los resultados puede afirmarse que los estudiantes son menos exigentes evaluando y puede

apreciarse en ellos una cierta falta de coherencia.

### Sobre el proceso: perspectiva émica

Comenzando por la primera gran familia de categorías, aspectos positivos a destacar, se han definido nueve categorías de análisis para clasificar la información procedente de la autoevaluación grupal del alumnado participante en el estudio. Sólo se valoraron unidades de análisis que tuvieran relación con los indicadores teórico-prácticos de evaluación para el aprendizaje. El porcentaje de fragmentos codificados en el análisis de contenido en cada categoría de análisis con respecto al total de fragmentos codificados en esta familia de categorías queda reflejado en la Tabla 4. Los conceptos a los que han hecho referencia con mayor frecuencia en sus textos los alumnos están relacionados con la distribución de tareas (262) y la asistencia y participación (230), aunque aspectos relacionados con valores (151) e ideas (111) también han sido elementos centrales de sus reflexiones.

Tabla 4.- Porcentaje de fragmentos del contenido de la autoevaluación codificados en cada categoría de análisis respecto a “aspectos positivos a destacar”.

Categorías de Análisis en <i>aspectos positivos a destacar</i>	%
Buena comunicación interna	1
Gestión del tiempo	3
Ideas aportadas al grupo	13
Asistencia y participación	27
Distribución de tareas	30
Práctica de contenidos de la asignatura	1
Recursos utilizados	5
Valor del papel del Profesor	2
Valor del aprendizaje cooperativo	18
Total	100

Aparecen cuestiones relacionadas con las aportaciones de los diferentes miembros de los grupos en lo que se refiere a la comunicación interna como fuente de aprendizaje cooperativo:

*‘Al igual que S1 y que S2, expresan mucho y se comunican muy bien dentro del*

*grupo, con buenas ideas para las actividades y con gran capacidad de aportar cosas muy interesantes, lo único que las cuesta más, hablar o expresarse en público, quizás sea por su edad.’*

Esta dimensión está íntimamente relacionada con las ideas aportadas al grupo, fuente de

conocimiento basada en experiencias personales:

*‘Aporta ideas relacionadas con la docencia, ya que tienen experiencias por familiares cercanos.’*

*‘Aporta experiencias en base a su hermano pequeño’.*

Aparecen referencias en cuanto a la gestión del conocimiento y toma de decisiones, relativas a la categoría anterior:

*‘Antes de realizar las prácticas hemos dado cada una opiniones e ideas para hacer una puesta en común y después hemos organizado el trabajo en función de la práctica.’*

*‘En general hemos aportado todas ideas y hemos consensuado con el grupo hasta llegar a un acuerdo final para completar las actividades’.*

También destacan las referencias en la categoría de asistencia y participación, recogiendo comentarios como:

*‘Las decisiones se han tomado teniendo en cuenta todas las aportaciones, realizando todos los trabajos*

*equitativamente, donde todos saliésemos beneficiados y aprendiéramos los contenidos de la asignatura, logrando siempre los objetivos propios de la tarea.’*

Aunque el porcentaje de representación sea mínimo conviene destacar el siguiente comentario en favor de la ayuda prestada por el docente al alumnado según la categoría valor del papel del profesor:

*‘Quizás hemos encontrado alguna dificultad a la hora de hacer alguna práctica, pero la profesora siempre ha estado a nuestra disposición para cualquier duda o problema ocurrido en las prácticas.’*

En relación con la familia de categorías dificultades encontradas, se han definido seis categorías para el análisis relacionadas con la evaluación para el aprendizaje. El porcentaje de fragmentos codificados en el análisis de contenido en cada categoría de análisis con respecto al total de fragmentos codificados en esta familia de categorías queda reflejado en la Tabla 5. En este caso, lo que más inquieta a los alumnos son las cuestiones relativas al tiempo (75) aunque diversos aspectos relacionados con las actividades (62) y la participación (42) también destacan entre sus preocupaciones.

Tabla 5.- Porcentaje de fragmentos del contenido de la autoevaluación codificados en cada categoría de análisis respecto a “dificultades encontradas”

Categorías de Análisis en dificultades encontradas	%
Contenidos trabajados en la asignatura	6
Gestión del tiempo	37
Recursos utilizados	4
Participación	21
Espacios en que se desarrollan las clases	1
Práctica de contenidos de la asignatura	31
Total	100

Dentro de las dificultades relacionadas con el desarrollo de las actividades prácticas realizadas a lo largo del curso –categoría de análisis s Práctica de contenidos de la asignatura– hay dos dimensiones que podemos

definir a partir de las aportaciones de los alumnos. La que aparece con más frecuencia es la referida a la utilidad de realizar actividades orientadas al análisis de casos reales, para establecer una conexión más

directa entre el conocimiento teórico y la práctica:

*‘Nos hubiera gustado haber realizado alguna práctica en una escuela infantil, aunque sabemos que no es posible.’*

*‘Habría sido más interesante poder realizarlo con un caso real para poder implicarnos más.’*

Esta necesidad que destacan se trata de solucionar a través de documentos o vídeos reales, pero dentro de los actuales planes de estudio no se contempla la posibilidad de realización de prácticas en centros durante el primer curso de grado en Magisterio. Una segunda dimensión es la que se orienta hacia el nivel de dificultad de las tareas así como la capacidad de los alumnos para abordarlas con éxito:

*‘Al principio no sabíamos cómo realizar el trabajo por sus dimensiones conceptuales, así como por la dificultad de éste.’*

Por último, hay múltiples referencias a cuestiones relacionadas con la organización del trabajo o la distribución del tiempo – categoría de análisis Gestión del tiempo– para realizar las actividades:

*‘Pensamos que con un poco más de tiempo y una corrección en clase habiéramos adquirido los conocimientos que la práctica nos quería aportar.’*

*‘Creemos que han sido demasiadas prácticas para tan poco tiempo, se nos juntaban unas con otras.’*

## **Discusión**

El objetivo de este trabajo era profundizar en el desarrollo de una experiencia de aprendizaje cooperativo, en especial en su proceso de evaluación de carácter participativo, para lo cual se utilizó como fuente de información los dos colectivos implicados en dicho proceso (profesorado vs alumnado).

Los resultados obtenidos del análisis de datos llevaron a concluir la apreciación de diferencias significativas entre las calificaciones otorgadas por el profesor y por el alumnado. La tendencia de los alumnos fue calificar los procesos de evaluación con puntuaciones más altas que el profesor. Tanto cuando realizaron autoevaluación como cuando realizaron evaluación entre iguales, en especial en esta última, se detectó escasa diferencia entre los diversos indicadores valorados en la evaluación entre iguales. Un análisis más detallado reflejó que las diferencias encontradas en los procesos de evaluación entre iguales fueron poco relevantes.

Los hallazgos encontrados deben observarse desde las restricciones metodológicas que este estudio tiene: muestra centrada en un único nivel de una única asignatura, titulación e institución académica; no permite realizar muchas generalizaciones, aunque tiene una amplia variabilidad dentro de dicho nivel de medida dado el tamaño de la muestra de alumnado participante en el estudio. En concordancia con lo que manifiesta Panadero, Alonso-Tapia y Huerta (2014) e investigaciones previas (Andrade, Du & Wang, 2008; Panadero & Jonsson, 2013; Sadler & Good, 2006), muestran que las Rúbricas bien diseñadas contribuyen a la mejora del rendimiento académico, aportan una diferenciación del mismo y enfatizan la evaluación del producto final. Según Panadero et al. (2014), este último aspecto “hace que los estudiantes presten mayor atención a sus calificaciones. [...] Debido a ello, las Rúbricas pueden hacer que los estudiantes se preocupen más de la puntuación y activen metas de rendimiento y/o evitación en lugar de metas de aprendizaje” (p. 166). Sin embargo, sus investigaciones no han dado resultados concluyentes. Para paliar dicho sesgo sobre su orientación al logro y valorar si en este estudio se dan resultados de evaluación para el aprendizaje se creó el cuestionario de autoevaluación grupal con preguntas abiertas cuyos resultados muestran evidencias que apuntan hacia la producción positiva de

aprendizaje. Por ello, podemos concluir que se ha dado evaluación para el aprendizaje.

Las limitaciones encontradas no descartaron el uso de la evaluación compartida como beneficiosa. Se comprobó que la evaluación compartida no es un sistema perfecto. No obstante, en el presente estudio también se probaron los beneficios y utilidad del proceso e instrumentos de evaluación creados según los resultados de los indicadores de la perspectiva ética del alumnado. El análisis de contenido de las reflexiones realizadas por el alumnado, que fueron recogidas con el instrumento de autoevaluación grupal, justifica esta afirmación. De tal forma que, al tratar de dar voz a los datos, se pudo oír como los estudiantes manifestaron: el desarrollo de procesos de comunicación positiva; enriquecimiento de los aprendizajes a través de las aportaciones de los miembros del grupo; desarrollo de habilidades sociales básicas que favorecen la autorregulación del trabajo grupal –distribución equitativa de la tarea– llegando a producirse “consenso”. Esto, en teoría, tiende a favorecer la cohesión grupal, sin embargo no se recogieron evidencias positivas al respecto. Es interesante la valoración positiva que realizan los estudiantes sobre la disponibilidad de la docente como recurso de apoyo. No obstante, también se recoge la evidencia empírica emergente en contra de los buenos resultados del proceso. Los estudiantes demandan sistemas de regulación externa del aprendizaje, en los que priman la atención del docente, la dedicación temporal a la realización de las prácticas en horario lectivo y una disminución del grado de dificultad de la tarea. Les resulta difícil autorregular su aprendizaje y buscan justificación para evitar llegar a este proceso. Lo más interesante de los resultados encontrados a este respecto, se centra en el inicio de dicho proceso –autorregulación del aprendizaje– en la mayoría de los estudiantes participantes en este estudio. Esto puede atribuirse, en parte, a la creación de situaciones de reflexión –individual vs grupal– que la docente propicia para la consecución de dicho logro durante los procesos de evaluación

compartida presente en este estudio. Por ello, reconsideramos la mejora de los diseños e instrumentos de evaluación como línea de investigación futura para propiciar un proceso de evaluación orientado al aprendizaje tanto de los estudiantes como del docente.

Si el objetivo de la evaluación universitaria no es valorar exclusivamente las competencias cognitivas, sino que debe considerar la valoración de otras competencias procedimentales y actitudinales, entendemos que la evaluación del aprendizaje realizada por el propio alumno constituye una variable especialmente relevante. Compartimos con Ibarra Saiz et al. (2012) que, a pesar de las dificultades, los estudiantes han logrado con esta experiencia: mejorar sus procesos y productos de aprendizaje; mejorar sus habilidades interpersonales; y, han experimentado las fortalezas y debilidades de una de las más complejas competencias profesionales del maestro –la evaluación–. La evaluación ha dejado de ser considerada una tarea periférica, realizada al final del proceso, para constituir una tarea de aprendizaje en sí misma, entendiendo que el alumno aprende reflexionando sobre su trabajo y el de sus compañeros (Gil Flores y Padilla, 2009).

## Referencias

- Alonso-Tapia, J. & Hernansaiz Garrido, H. (2013). Evaluar para el aprendizaje. Problemas en la práctica de la evaluación de competencias: el caso de la comprensión de tablas y gráficos. *Assessment for learning. Difficulties in competencies practices: the case of table and graph understanding. Revista educativa Hekademos*, 13, Año VI, 7-18.
- Andrade, H., Du, Y. & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practices*, 27(2), 3-13.

- Boud, D. & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31 (4), 399-413. Disponible en [http://pages.ramapo.edu/~vasishth/Learning\\_Outcomes/Boud+Assessment+Long-term\\_Learning.pdf](http://pages.ramapo.edu/~vasishth/Learning_Outcomes/Boud+Assessment+Long-term_Learning.pdf)
- Bretones Román, A. (2008). Participación del alumnado de Educación Superior en su evaluación. *Revista de Educación*, 347, pp.181-202.
- Brown, S. & Glasner, A. (2003). *Evaluar en la Universidad. Problemas y nuevos enfoques*. Madrid: Narcea.
- Carless, D., Joughin, G. & Liu, N.F. (2006). *How assessment supports learning: Learning-oriented assessment in action*. Hong Kong University-Press.
- Cochran-Smith, M. & Lytle, S.L. (2002). *Dentro / Fuera. Enseñantes que investigan*. Madrid: Akal.
- Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self-assessments. *Assessment & Evaluation in Higher Education*, 11 (2), 144-166.
- Falchikov, N. y Boud, D. (1989). Student self-assessment in Higher Education: a meta-analysis. *Review of Educational Research*, 59 (4), pp.395-430.
- Falchikov, N. (1995). Peer feedback marking: developing peer assessment. *Innovations in Education and Training International*, 32, 175-187.
- Falchikov, N. (2005). *Improving Assessment Through Student Involvement. Practical solutions for aiding learning in higher and further education*. London: Routledge-Falmer.
- Gallego, B.; Sánchez, C. & Cubero, J. (2011). La pro-alimentación y la retro-alimentación prospectiva entre iguales como estrategias para el desarrollo del pensamiento crítico y el aprendizaje autónomo. En Evalfor, *Evaluar para aprender en la universidad: Experiencias innovadoras en el aprendizaje a través de la evaluación* (pp. 83-97). Madrid: Bubok Publishing.
- Gil Flores, J. & Padilla, T. (2009). La participación del alumnado universitario en la evaluación del aprendizaje. *Educación XXI*, 12(1), pp. 43-65.
- Heathfield, M. (2003). Evaluación en grupo para fomentar un aprendizaje de calidad. En S. Brown, A. Glasner (Eds.) *Evaluar en la universidad. Problemas y nuevos enfoques*, (pp.155-166). Madrid: Narcea.
- Ibarra Saiz, M.S., Rodríguez Gómez, G. & Gómez Ruíz, M.A. (2012). La evaluación entre iguales: beneficios y estrategias para su práctica en la universidad. *Revista de Educación*, 359, 206-231.
- López Pastor, V.M. (2008). Desarrollando sistemas de evaluación formativa y compartida en la docencia universitaria. Análisis de resultados de su puesta en práctica en la formación inicial del profesorado. *European Journal of Teacher Education*, 31(3), 293-311.
- López Pastor, V.M. (Coord.) (2009). *Evaluación formativa y compartida en Educación Superior. Propuestas, técnicas, instrumentos y experiencias*. Madrid: Narcea.
- Montero-Fleta, B. (2006). Investigación cualitativa y cuantitativa en la coevaluación: correlación profesor-alumno. En F. Watts & A. García-Carbonell, (Eds.), *La evaluación compartida: investigación multidisciplinar*, (pp. 31-46). Universidad Politécnica de Valencia. Recuperado de <http://www.upv.es/gie/Publicaciones.html>
- Morales Vallejo, P. (2008). *Estadística Aplicada a las Ciencias Sociales*. Madrid: Universidad Pontificia de Comillas.
- Muñiz, J. & Fonseca-Pedrero, E. (2008). Construcción de instrumentos de medida para la evaluación universitaria. *Revista de Investigación en Educación*, 5, 13-25.

- Navaridas Nalda, F. (2002). La evaluación del aprendizaje y su influencia en el comportamiento estratégico del estudiante universitario. *Contextos Educativos*, 5, 141-156.
- Otero, M., Yuste, R. & Alzás, T. (2011). Coevaluación en la plataforma virtual Moodle: una propuesta práctica para la mejora del proceso de enseñanza en Educación Superior. En Evalfor, *Evaluar para aprender en la universidad: Experiencias innovadoras en el aprendizaje a través de la evaluación*, (pp. 62-71). Madrid: Bubok Publishing.
- Panadero, E. & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9(0), 129-144. DOI: [10.1016/j.edurev.2013.01.002](https://doi.org/10.1016/j.edurev.2013.01.002)
- Panadero, E. & Alonso-Tapia, J. (2013). Autoevaluación: Connotaciones teóricas y prácticas. Cuándo ocurre, cómo se adquiere y qué hacer para potenciarla en nuestro alumnado. *Electronic Journal of Research in Educational Psychology*, 11 (2), 551-576. Disponible en: <http://www.redalyc.org/articulo.oa?id=293128257012>
- Panadero, E., Romero, M., & Strijbos, J. W. (2013). The impact of a rubric and friendship on construct validity of peer assessment, perceived fairness and comfort, and performance. *Studies in Educational Evaluation*, 39(4), 195-203. DOI: [10.1016/j.stueduc.2013.10.005](https://doi.org/10.1016/j.stueduc.2013.10.005)
- Panadero, E., Alonso-Tapia, J. & Huertas, J.A. (2014). Rubrics vs self-assessment scripts: effects on first year university students' self-regulation and performance / Rúbricas y guiones de autoevaluación: efectos sobre la autorregulación y el rendimiento de estudiantes universitarios de primer año. *Infancia y Aprendizaje: Journal for the Study of Education and Development*, 37(1), 149-183. DOI: [10.1080/02103702.2014.88165555](https://doi.org/10.1080/02103702.2014.88165555)
- Pond, K., Ul-Haq, R. & Wade, W. (1995). Peer review: a precursor to peer assessment. *Innovation in Education and Training International*, 32, 314-323.
- Roberts, T. S. (2006). *Self, Peer and Group Assessment in E-Learning*. Hershey, PA: Information Science Publishing.
- Rodríguez Gómez, G. & Ibarra Saiz, M.S. (Eds.) (2011). *e-Evaluación orientada al e-Aprendizaje estratégico en Educación Superior*. Madrid: Narcea.
- Rodríguez Gómez, G., Ibarra Sáiz, M.S. & García Jiménez, E. (2013). Autoevaluación, evaluación entre iguales y coevaluación: conceptualización y práctica en las universidades españolas. *Revista de Investigación en Educación*, 11 (2), 198-210. Recuperado de <http://dialnet.unirioja.es/servlet/articulo?codigo=4734976>
- Sadler, P.M. & Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment*, 11(1), pp. 1-31.
- Sanmartí, N. (2007). *Evaluar para aprender: 10 ideas clave*. Barcelona: Graó.
- Topping, K.J. (2009). Peer assessment. *Theory Into Practice*, 48, 20-27.

# ANEXO 1

## Instrumentos utilizados

### *Rúbrica de Evaluación entre iguales*

	Alumno1	Alumno2	Alumno 3
Asiste regularmente a las reuniones del grupo			
Aporta ideas para el trabajo			
Trae material de investigación, analiza y prepara para el tema			
Contribuye a los procesos cooperativos del grupo			
Apoya y motiva a los miembros del grupo			
Contribuye al resultado final y a la presentación			
Observaciones			

10	5	0
<b>Asistencia a las reuniones del grupo</b>		
Asiste a todas las reuniones, presenciales o virtuales permaneciendo hasta el final, trabajando dentro del tiempo de forma activa y atenta, preparado para la flexibilidad de horarios.	Se perdió varias reuniones, demuestra desinterés por el trabajo en grupo	Se perdió la mayoría de las reuniones. Siempre o a menudo llega tarde, no atiende o pierde la mayor parte del tiempo.
<b>Aportación de ideas para el trabajo</b>		
Piensa sobre el tema antes de la reunión, proporcionando ideas para trabajar, trabaja sobre las sugerencias de otros y está preparado para rebatir las ideas más que para quedarse callado.	No viene preparado. Contribuye con alguna idea. Ni rechaza ni debate las ideas.	No viene preparado. No contribuye con ninguna idea. Tiende a rechazar las ideas de los otros en lugar de rebatirlas.
<b>Material de investigación , análisis y preparación del tema</b>		
Hace lo que se le dice que haga, trae materiales, hace un reparto equitativo de la investigación y ayudó a analizar y valorar el material.	No siempre hace lo que se le dice que haga, no aporta materiales y no hace un reparto igualitario de las tareas, suele salir beneficiado en dicho reparto.	No investiga. No hace lo que promete hacer. No gestiona su carga de trabajo. No se implica en las tareas y permite que los demás proporcionen todo el material.
<b>Contribución a los procesos cooperativos del grupo</b>		
Deja las diferencias personales fuera del grupo, desea revisar el progreso y evita el conflicto, adopta los distintos roles necesarios, mantiene al grupo por el buen camino, es flexible pero se centra en el tema.	Deja las diferencias personales fuera del grupo. Siempre adopta el mismo rol (líder, bromista, etc.) a pesar de las circunstancias. No toma la iniciativa para revisar y progresar en el trabajo.	No toma la iniciativa, esperando a que se le diga qué ha de hacer. Siempre adopta el mismo rol (líder, bromista, etc.) a pesar de las circunstancias, crea el conflicto y no está preparado para revisar el progreso del grupo.

<b>Apoyo y motivación a los miembros del grupo</b>		
Bueno escuchando a los otros, motiva la participación, se adapta a un ambiente cooperativo de aprendizaje, sensible a los temas que afectan a los miembros del grupo, apoya a los miembros del grupo con necesidades especiales.	Solo se preocupa de completar la tarea; aunque escucha a los demás no tiene en cuenta sus opiniones al realizar el trabajo, no motiva ni apoya en la participación de los compañeros.	Solo se preocupa de completar la tarea, habla sobre los otros e ignora sus opiniones. Insensible a las necesidades individuales y no contribuye al proceso de aprendizaje.
<b>Contribución práctica al resultado final, presentación</b>		
Desea intentar hacer cosas nuevas. No estropea las tareas, hace una contribución de alto nivel, toma su propia iniciativa, creíble y produce un trabajo/ presentación de alta calidad.	Se compromete con algunas tareas, haciendo una contribución de nivel medio. No toma su propia iniciativa pero realiza un trabajo de buena calidad, haciendo lo que se le dice que haga.	No desea comprometerse con ninguna tarea, no acepta responsabilidades, no es creíble. Contribuye de una forma pobre y limitada.

(Adaptado de Heathfield, 2003, p. 162)

### **Cuestionario de Autoevaluación**

Este cuestionario pretende indagar sobre cómo ha tenido lugar la experiencia de trabajo en grupo. Consta de 4 bloques en los que tenéis que reflexionar sobre:

- a. Los objetivos comunes logrados por todo el grupo. Hace referencia a las dificultades encontradas a lo largo del desarrollo del trabajo y cómo han ido encontrándose soluciones.
- b. Los acuerdos o decisiones que se han ido tomando, referido principalmente a las decisiones como grupo.
- c. Las aportaciones que individualmente ha realizado cada miembro del grupo
- d. Una calificación grupal negociada (en base 10)

<b>Participantes</b>	<b>Responsabilidades</b>
<b>Bloque 1: Dificultades encontradas</b>	
<b>Bloque 2: Acuerdos y decisiones</b>	
<b>Bloque 3: Aportaciones</b>	
Miembro del grupo	Aportación realizada
<b>Bloque 4: Calificación Grupal (negociada)</b>	
<b>Otras Observaciones</b>	

### **Escala de valoración del profesor**

<b>Actividades entregadas</b>							
Nivel 1 Muy Bueno		Nivel 2 Bueno		Nivel 3 Apto		Nivel 4 Repetir	
<b>Organización, desarrollo y claridad de las actividades</b>							
Nivel 1 Muy Bueno		Nivel 2 Bueno		Nivel 3 Apto		Nivel 4 Repetir	
<b>Reflexiones sobre la aportación de la práctica y conexión con la teoría</b>							
Nivel 1 Muy Bueno		Nivel 2 Bueno		Nivel 3 Apto		Nivel 4 Repetir	
<b>Consultas bibliográficas realizadas</b>							
Nivel 1 Muy Bueno		Nivel 2 Bueno		Nivel 3 Apto		Nivel 4 Repetir	
<b>Aspectos formales</b>							
Nivel 1 Muy Bueno		Nivel 2 Bueno		Nivel 3 Apto		Nivel 4 Repetir	

#### **Autores**

**To know more / Saber más**

**Pascual-Gómez, Isabel** ([isabel.pascualg@uah.es](mailto:isabel.pascualg@uah.es)).

Doctora en Filosofía y Ciencias de la Educación por la Universidad Complutense de Madrid. Actualmente, profesora contratada-doctor en Métodos de Investigación y Diagnóstico en Educación del Departamento de Ciencias de la Educación de la Universidad de Alcalá. Entre sus líneas de investigación destacan la evaluación de la docencia universitaria y la validación de instrumentos psicopedagógicos.



**Lorenzo-Llamas, Elena María** ([elenam.lorenzo@uah.es](mailto:elenam.lorenzo@uah.es)).

Doctora en Psicopedagogía por la Universidad de Alcalá. Actualmente, profesora-asociada en Métodos de Investigación y Diagnóstico en Educación del Departamento de Ciencias de la Educación de la Universidad de Alcalá. Entre sus líneas de investigación destacan las altas capacidades intelectuales, métodos de investigación en educación, atención a la diversidad e inclusión y sistemas de evaluación en Educación Superior.



**Monge-López, Carlos** ([carlos.monge@edu.uah.es](mailto:carlos.monge@edu.uah.es)).

Doctorando en Planificación e innovación educativa. Personal investigador del Departamento de Ciencias de la Educación, área de Didáctica y Organización Escolar de la Universidad de Alcalá. Ostenta un contrato Pre-doctoral de Formación de Doctores FPU-UAH.





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

[ISSN: 1134-4032]

© Copyright, RELIEVE. Reproduction and distribution of this articles it is authorized if the content is no modified and their origin is indicated (RELIEVE Journal, volume, number and electronic address of the document).

© Copyright, RELIEVE. Se autoriza la reproducción y distribución de este artículo siempre que no se modifique el contenido y se indique su origen (RELIEVE, volumen, número y dirección electrónica del documento).