

PROCESO DE EVALUACIÓN ENTRE PARES EN LA RESOLUCIÓN DE PROBLEMAS EN FUTUROS DOCENTES. ESTUDIO DE UN CASO

Peer assessment process in solving problem with prospective teachers. Case study

de Armas-González, P., Perdomo-Díaz, J. y Sosa-Martín, D. N.

Universidad de La Laguna

Resumen

Se presenta parte de una investigación cuyo objetivo es analizar el proceso que siguen docentes en formación al evaluar una actividad matemática realizada por un par. Los participantes son estudiantes del Máster de Formación del Profesorado de Educación Secundaria. Se les pidió resolver un problema matemático y, por parejas, evaluar la resolución de otro compañero. En este documento mostramos el estudio de un caso, correspondiente al proceso de evaluación realizado por una de las parejas. Se analizan las resoluciones de los dos evaluadores y del evaluado y la transcripción de la discusión mantenida por la pareja mientras evaluaba, identificando aquellos aspectos en los que se fijaban y el carácter de sus comentarios. Esta pareja se centra en el proceso general de resolución y sus reflexiones son tanto interpretativas como valorativas. Además, muestran dificultades para evaluar a su compañero y un fuerte apego a su propia resolución.

Palabras clave: *evaluación entre pares, resolución de problemas, formación docente*

Abstract

Part of an investigation is presented whose objective is to analyse the process that teachers in training follow when evaluating a mathematical activity carried out by a peer. Participants are students of the Secondary Education Teacher Training Master. They were asked to solve a mathematical problem and, in pairs, to evaluate the resolution of another classmate. In this document we show the case study, corresponding to the evaluation process carried out by one of the couples. It is analysed the resolutions of the two evaluators and of the evaluated and the transcript of the discussion held by the couple while evaluating, identifying those aspects in which they focused and the nature of their comments. This couple focuses on the general resolution process and their reflections are both interpretive and evaluative. Furthermore, they show difficulties in evaluating their partner and a strong attachment to their own resolution.

Keywords: *peer assessment, problem solving, teacher training*

INTRODUCCIÓN

La evaluación es una herramienta poderosa que se usa para una gran variedad de propósitos como llevar un control del aprendizaje de los alumnos, evaluar sus logros, tomar decisiones para la enseñanza, evaluar el currículo, entre otros (Silver y Mills, 2018). Durante la última década se ha intentado motivar a los profesores para que incorporen en sus clases prácticas de evaluación que vayan más allá del tradicional examen escrito, a la par que se investiga sobre el efecto de incorporar estos métodos alternativos de evaluación en la práctica docente. La evaluación entre pares es una de las estrategias que mayor presencia está cobrando en este escenario, junto con la autoevaluación. Distintas investigaciones muestran que la evaluación entre pares aporta beneficios al aprendizaje (Topping, 2009), contribuyendo a aumentar tanto el conocimiento de conceptos como de

herramientas o estrategias para enfrentarse a diversas situaciones, y a reconocer argumentos válidos (Beaver y Beaver, 2011; Lavy y Shriki, 2014). Custodia, Márquez y Sanmartí (2015) muestran que este enriquecimiento se da tanto en el sujeto evaluador como en su par evaluado.

A pesar de los beneficios que ha mostrado el uso de estrategias de evaluación que involucren a los propios estudiantes, son muy pocos los docentes que implementan estos métodos en el aula (Chanudet, 2016). Más aún si nos fijamos en la evaluación por pares, como muestra el trabajo de Suurtam, Koch y Arden (2010), en el que 1096 profesores de 7º y 10º grado de escuelas de Ontario, Canadá, respondieron a un cuestionario. El 91% de los profesores admitió sentirse cómodo usando distintos tipos de evaluación, siendo los más usados los exámenes escritos (95%) y los cuestionarios (84%). La evaluación por pares solo la utilizaba un 15%, el 10% como herramienta para hacerse una idea del aprendizaje los alumnos y solo un 5% como estrategia para evaluar.

En las investigaciones de Beaver y Beaver (2011) y Lavy y Shriki (2014) son futuros profesores de matemáticas y profesores en ejercicio los que llevan a cabo una actividad de evaluación entre pares. Los resultados mostraron que este tipo de evaluación contribuye al desarrollo de los docentes en diversos aspectos: (i) mejora su habilidad para evaluar, ya que eligen mejor los criterios y el peso numérico que se asignan a cada criterio, (ii) aumenta su propio conocimiento sobre los contenidos matemáticos y las herramientas con las que resuelven problemas, (iii) les ayuda a darse cuenta de que hay muchas formas diferentes de resolver un mismo problema, haciéndoles reflexionar sobre su propia solución y sobre cómo mejorarla, y (iv) les proporciona experiencia en el reconocimiento de argumentos válidos, lo que les conduce a ganar confianza para ser críticos con el razonamiento de los demás y, en este sentido, a ser responsables con los comentarios que brindan como evaluadores.

Los dos trabajos anteriores tienen en común que los participantes realizaron la evaluación entre pares a partir de una serie de criterios predefinidos. En el caso de Beaver y Beaver (2011), se les entregó una rúbrica con la que tenían que evaluar las producciones de sus compañeros; mientras que Lavy y Shriki (2014) pidieron a los participantes que construyeran, de forma conjunta, la lista de criterios que pretendían evaluar. En ambos estudios, las evaluaciones entre pares se hicieron a partir de unas referencias comunes para todos los evaluadores. Esas referencias, en forma de criterios de evaluación o rúbricas, son complejas de definir, conllevan un proceso de construcción que implica reflexión, puesta en común y toma de acuerdos. En numerosas ocasiones, cuando los docentes tienen que evaluar una actividad, carecen de este tipo de herramientas. ¿Qué hacen en esos casos? ¿Cómo definen sus criterios de evaluación? ¿Qué toman en cuenta? Estas cuestiones cobran mayor relevancia en el ámbito de la formación inicial docente, que es donde se sitúa esta investigación.

La pregunta general que guía esta investigación es: ¿Qué caracteriza al proceso que siguen los docentes de matemáticas de secundaria en formación cuando se les pide evaluar una actividad matemática realizada por un par? Nos planteamos entonces los siguientes objetivos:

- Identificar qué aspectos de la resolución de la actividad matemática tienen en cuenta los futuros docentes cuando actúan como evaluadores.
- Analizar el carácter de sus reflexiones.

Para tratar de abordar estos objetivos se pidió a los estudiantes de la especialidad de Matemáticas del Máster en Formación del Profesorado de Educación Secundaria de la Universidad de La Laguna, que resolvieran individualmente un problema matemático. Posteriormente, en parejas, debían evaluar la forma en que uno de sus compañeros había resuelto la misma actividad.

En total participaron 16 estudiantes. En este trabajo se presenta el análisis de un caso, formado por una pareja de estudiantes de dicho Máster. Se analiza, por un lado, la forma en que el estudiante evaluado y los evaluadores resolvieron el problema, y por otro lado, la discusión que la pareja de evaluadores mantuvo durante la evaluación de la respuesta de su compañero.

Arnal-Bailera, Muñoz-Escolano y Oller-Marcén (2018) y Cárdenas, Blanco, Guerrero y Caballero (2016) identifican algunos aspectos que los profesores tienen en cuenta cuando califican producciones de estudiantes, entre los que se incluyen el proceso general de resolución, los cálculos realizados y el resultado obtenido. Estos estudios fueron realizados con profesores en ejercicio, en un contexto de heteroevaluación. La investigación que aquí se presenta mostrará si esos mismos aspectos son los que tienen en cuenta los docentes en formación, en un contexto de evaluación entre pares, o si aparecen nuevos elementos.

MARCO CONCEPTUAL

En numerosas ocasiones la evaluación se asocia al hecho de asignar una calificación. Sin embargo, el papel de la evaluación en el aula va más allá (Beaver y Beaver, 2011). Actualmente, la visión de la evaluación como una serie de métodos que miden objetivamente la adquisición de conocimientos ha cambiado hacia una visión de la evaluación como una práctica que proporciona conocimientos e información de forma continua, que apoya el aprendizaje del alumno e influye en la práctica docente (Surtamm, Koch y Arden, 2010).

Shahbari y Abu-Alhija (2018) definen la evaluación como el proceso de recopilar información sobre el aprendizaje de los alumnos, interpretar esa información y llevar a cabo una acción. Nortvedt, Santos y Pinto (2016) añaden que la información recogida puede ser usada tanto por profesores como por estudiantes, y que el fin de la evaluación es identificar dónde se encuentran los alumnos en su aprendizaje, a dónde deben dirigirse y cuál es la mejor forma de hacerlo.

La evaluación se puede llevar a cabo de varias maneras y puede ser usada con diferentes propósitos. Uno de los métodos que se puede emplear es la evaluación entre pares, definida por Topping (2009) como un método en el que los estudiantes determinan el nivel o la calidad de una tarea o del rendimiento de otros estudiantes de su mismo nivel. Las actividades de evaluación por pares pueden ser variadas, funcionando en diferentes áreas del currículo y en diferentes asignaturas. El producto a evaluar también puede variar: redacciones, portafolios, presentaciones orales, pruebas escritas, etc. Algunos autores sostienen la necesidad de proporcionar criterios de evaluación a la hora de llevar a cabo una actividad de evaluación por pares (Beaver y Beaver, 2011; Custodia, Márquez y Sanmartí, 2015; Lavy y Shriki, 2014). Sin embargo, otros estudios muestran que cuando se evalúa en grupo, entregar materiales para que evalúen no necesariamente ayuda a llegar a un consenso sobre cómo evaluar o, ni siquiera, sobre cómo utilizar esos materiales (Wyatt-Smith, Klenowski y Gunn, 2010).

En este estudio, la evaluación por pares se entenderá como el proceso que siguen los estudiantes, en este caso futuros profesores, para analizar la respuesta de un compañero a una tarea determinada, en este caso la resolución a un problema matemático. En este proceso pueden intervenir diversos elementos, entre ellos el hecho de que la persona que evalúa podría tener que analizar una respuesta completamente diferente a la suya, lo que pondría en juego sus propios conocimientos acerca de las diferentes formas de resolver la actividad (Cáceres y Chamoso, 2015; Cárdenas et al., 2016).

Por otra parte, cada evaluador puede centrarse en el análisis de diferentes aspectos de la resolución presentada por su par. En el caso particular de la evaluación de actividades matemáticas, Arnal-Bailera, Muñoz-Escolano y Oller-Marcén (2018) identifican cinco aspectos: (i) procedimiento, que incluye referencias al proceso general de resolución, (ii) cálculos, que incluye cualquier tarea concreta dentro del proceso general de resolución, (iii) errores, (iv) exposición, haciendo referencia a la claridad en la exposición de la respuesta y (v) resultado. Las reflexiones en torno a esos aspectos pueden tener, además, carácter interpretativo o valorativo.

METODOLOGÍA

Esta investigación es de tipo cualitativo, con un enfoque descriptivo exploratorio. Se realiza el estudio de un caso, con el objeto de analizar en detalle las variables deseadas (Creswell, 2012).

Participantes

La población objeto de estudio eran 16 estudiantes de la especialidad de Matemáticas del Máster en Formación del Profesorado de Educación Secundaria Obligatoria y Bachillerato de la Universidad de La Laguna que estaban cursando la asignatura “Aprendizaje y enseñanza de las Matemáticas”. En este trabajo se presenta el estudio de un caso en el que están involucrados 3 de esos estudiantes. Nos referiremos a ellos con los pseudónimos Carlos (evaluado) y Ana y Pedro (evaluadores). Los dos primeros son graduados en Matemáticas y el tercero es Licenciado en Ciencias Estadísticas. Se eligió analizar este caso, de forma intencional, al observarse que la pareja de evaluadores era una de la parejas que más debatía y que, por tanto, podrían aportar mayor riqueza al análisis.

Recogida de datos

Los datos se recogieron a partir de una actividad, realizada durante una sesión de 2 horas de clase, que incluía cuatro fases: resolución de un problema, evaluación entre pares, autoevaluación y discusión final. Nos centraremos, en este trabajo, en el análisis de las dos primeras fases.

La primera fase consistió en una actividad que tenían que resolver de forma individual, con lápiz y papel (Figura 1) y que se recogió una vez finalizaron su resolución. El criterio que se utilizó para seleccionar la actividad fue que resultara ser un problema para los participantes, en el sentido de Schoenfeld (1985). En la segunda fase se formaron aleatoriamente ocho parejas y a cada pareja se le entregó la resolución de un compañero, junto con un documento con las indicaciones para la evaluación (Figura 2). Esta fase se grabó en audio con el objetivo de obtener la mayor información posible del proceso que seguían al evaluar.

Encontrar ecuaciones de rectas que tengan dos puntos de intersección con la parábola $y = x^2 + 4x + 5$.

Figura 1. Problema propuesto a los estudiantes.

Evalúa la resolución del problema que se te entrega en documento aparte. Argumenta de forma detallada dicha evaluación.

(No escribir sobre el documento de resolución del problema)

Figura 2. Indicaciones para la evaluación por pares.

Proceso de análisis de datos

Los instrumentos utilizados para el análisis fueron: la resolución del problema entregada por los tres estudiantes y el archivo de audio con la discusión mantenida por la pareja evaluadora.

El proceso de análisis se realizó en dos fases. En primer lugar, se analizaron las resoluciones individuales de Ana, Pedro y Carlos al problema planteado, identificando los heurísticos utilizados, los casos que estudiaron, el proceso seguido, las representaciones que emplearon y las soluciones dadas. Posteriormente se analizó el proceso que siguieron Ana y Pedro para evaluar la resolución de Carlos, a partir de la transcripción del audio, teniendo en cuenta tres variables: resolución a la que se refieren (Carlos, Ana y/o Pedro), aspecto en el que se fijan y carácter del comentario (expositivo, interpretativo o valorativo). Para establecer los valores de la segunda variable, se tomaron como referencia las categorías definidas por Arnal-Bailera, Muñoz-Escolano y Oller-Marcén (2018) que aparecen en el marco y que fueron adaptadas a las características de la actividad que aquí se evaluaba. Se consideraron por tanto las siguientes categorías, atendiendo al aspecto de la resolución del problema en que se podían fijar los evaluadores:

- Proceso general: Comentarios acerca del proceso general seguido en la resolución.
- Operatoria: Comentarios de una tarea concreta dentro del proceso general de resolución.

- Argumentación: Comentarios acerca de las explicaciones dadas por el evaluado.
- Resultado: Comentarios acerca del resultado final presentado por el alumno.

Para sistematizar el análisis se transcribió el audio con la discusión de Ana y Pedro. La transcripción se dividió en fragmentos que denominamos episodios, esta división se realizó atendiendo a la categoría en la que podían enmarcarse los comentarios que los evaluadores hacían en dicho fragmento (proceso general, operatoria, argumentación o resultado). A su vez, dentro de cada episodio, se realizó una nueva división a cuyos fragmentos denominamos sub-episodios. Estos sub-episodios se obtuvieron atendiendo al carácter de dichos comentarios, es decir, si los comentarios eran de carácter interpretativo, valorativo o expositivo.

ANÁLISIS Y RESULTADOS

La Tabla 1 muestra los heurísticos empleados por Ana, Pedro y Carlos para resolver el problema.

Tabla 1. Heurísticos empleados durante la resolución del problema.

Evaluadores		Evaluado
Ana	Pedro	Carlos
Representa gráficamente la parábola. Se basa en dicha representación para resolver el problema.	Divide en casos. Primero busca rectas horizontales, haciendo un estudio algebraico, y continúa con otros casos.	Estudia las expresiones algebraicas de la parábola y las rectas.

Más detalladamente, en la resolución de Ana observamos los siguientes pasos: (1) Calcula los puntos de corte de la parábola con los ejes; (2) Representa gráficamente la parábola (comete un error al representar el vértice); (3) Indica dos soluciones particulares de rectas horizontales y, para verificarlo, calcula los puntos de corte de las dos rectas con la parábola; (4) Generaliza (sin justificar) para rectas horizontales por encima de “su vértice”.

Pedro divide su resolución en casos en los que realiza los pasos indicados a continuación:

- Caso 1. Busca rectas horizontales que corten a la parábola en dos puntos: (1) Calcula la recta tangente a la parábola que tiene pendiente 0, mediante la derivada. (2) Generaliza (sin justificar) para rectas horizontales por encima de la anterior.
- Caso 2. Busca rectas que pasen por el vértice y por cualquier otro punto de la parábola. Sin dar la expresión algebraica, indica que es todo el haz de rectas que pasan por el vértice a excepción de la horizontal y la vertical.
- Caso 3. Indica, sin realizar ningún cálculo, que intentaría generalizar a rectas que pasen por dos puntos arbitrarios de la parábola.

Por último, Carlos sigue los siguientes pasos: (1) Iguala la ecuación de la parábola a la ecuación general explícita de una recta ($y=ax+b$) y pasa todos los términos a un miembro (comete errores de cálculo); (2) Resuelve la ecuación de 2º grado, e indica que el discriminante de la raíz debe ser estrictamente positivo; (3) Plantea la inequación resultante y comienza a estudiarla (comete errores de cálculo); (4) Resuelve la ecuación considerando la igualdad y calcula el valor de a respecto de b ; (5) Resuelve la inequación y calcula los intervalos de los posibles valores de a si $b=0$.

Análisis de la evaluación entre pares

Del análisis de la transcripción se obtuvieron un total de 24 episodios y 43 sub-episodios. La Tabla 2 muestra el número de episodios asociado a cada categoría y el carácter interpretativo (I), valorativo (V) o expositivo (E) de los sub-episodios.

Los evaluadores se centraron mayoritariamente en aspectos relacionados con el proceso general de la resolución del problema. De ahí que, aun siendo 7 episodios los que aparecen en esta categoría,

se dividan en un total de 24 sub-episodios. En 3 de esos 7 episodios, la discusión gira en torno a las tres soluciones individuales, la del evaluado y las de los evaluadores. Son episodios donde los evaluadores comparan sus propias resoluciones con la del evaluado, como el siguiente:

1. P: O sea, tú y yo, la casualidad que hicimos más o menos el mismo razonamiento.
2. P: Yo lo que digo es, lo que está haciendo al estar con $y=ax+b$ y de repente decir b es igual cero, es que está cogiendo todas las rectas que pasan por el $(0,0)$. Porque se queda con $y=ax$.
3. P: La parábola tiene, tú la calculaste más o menos igual que yo, tiene una forma más o menos así.

Tabla 2. Número de episodios en cada categoría y carácter de los mismos.

	N° total de episodios (sub episodios)	Resolución a la que se refieren			Carácter		
		Evaluado (Carlos)	Evaluador 1 (Ana)	Evaluador 2 (Pedro)	I	V	E
Proceso general	7 (24)	6	4	3	8	11	5
Operatoria	4 (5)	4	0	1	3	2	0
Argumentación	1 (1)	1	0	0	1	0	0
Resultado	5 (6)	5	0	1	2	4	0
Otros	7 (7)	-	-	-	-	-	-
Total		16	4	5	14	17	5

Algunos de esos episodios ponen de manifiesto las dificultades con las que se encuentra la pareja de evaluadores para poder comprender la resolución de su compañero. Por otro lado, esos episodios también muestran cierto apego de los evaluadores hacia sus propias resoluciones, tomándolas en varios momentos como referente para analizar la resolución de su compañero.

Estas dificultades que tienen los evaluadores para entender la resolución de Carlos también pueden verse en los episodios clasificados en las categorías de Operatoria y Resultado. Durante esos 4 y 5 episodios, respectivamente, Ana y Pedro continúan en la línea de intentar comprender los pasos seguidos por su compañero.

Además, el único comentario clasificado en la categoría de Argumentación también está relacionado con dichas dificultades, un episodio en el que Pedro explica que echa en falta algún tipo de argumentación con palabras en la resolución de Carlos.

En el análisis se identificaron siete episodios que no pudieron clasificarse en ninguna de las cuatro categorías definidas en el marco (Tabla 2). En estos episodios, los evaluadores hacen referencia al tipo de problema que se les planteó [4], sus posibles soluciones [5-6], su autopercepción como resolutores de problemas [7] y sobre el proceso de evaluación entre pares [8-9].

4. P: Yo estoy bastante oxidado con esto. Yo pensé que iba a ser resolver problemas como los de secundaria. Pero esto ya se me queda largo.
5. A: Y, en verdad, no te dice que encuentres todas. Te dice que encuentres ecuaciones de la recta.
6. P: Ecuaciones de rectas. Con dos ya sería suficiente. Fue lo primero que pensé
7. P: Yo es que no me sentía capaz tampoco de resolverlo de otra forma, pero claro.
8. A: Pero ¿hay que ponerle una nota o algo?
9. P: No, no, no. No es corregir, ¿no? Bueno, en principio yo no lo evaluaría porque tampoco sé sobre qué lo tengo que evaluar, si hay que ponerle una nota, si hay que poner más o menos iba bien.

DISCUSIÓN FINAL

En este trabajo se muestra el primer paso de un estudio que pretende dar respuesta a la siguiente pregunta de investigación: ¿qué caracteriza al proceso que siguen los docentes de matemáticas de secundaria en formación cuando se les pide evaluar una actividad matemática realizada por un par?

Del estudio de este caso se desprende que una de las características de ese proceso tiene que ver con el grado de apego que los evaluadores tengan a su propia resolución. Esto se pudo observar gracias al análisis previo que se hizo de las resoluciones de los tres estudiantes. En más de la tercera parte de los episodios, los evaluadores hacen referencia a sus propias estrategias de resolución (Tabla 2), tratando de comprobar si sus soluciones están contempladas en la resolución que están evaluando. Este apego, en algunos casos, puede influir en el proceso de evaluación (Cárdenas et al., 2016).

Otro de los resultados relevantes de este análisis, y que podría explicar esta dependencia de su solución, es la evidente dificultad que manifiestan para poder comprender la resolución a evaluar. Aunque, a excepción de la argumentación, la pareja analiza todos los aspectos de la resolución considerados en las categorías, este análisis no sigue un orden lógico. No analizan todos los pasos de la resolución y, los que analizan, lo hacen sin ser capaces de conectar los argumentos utilizados por su compañero. Durante el proceso de evaluación, la pareja realiza comentarios tanto de carácter interpretativo como valorativo (Tabla 2). Existen episodios en los que al realizar comentarios de carácter valorativo, atribuyen al compañero evaluado errores que no comete y que son fruto de la falta de comprensión a la hora de interpretar ciertos pasos de la resolución. Y, sin embargo, pasan por alto errores en la operatoria que sí comete su compañero. Con esto se observa como uno de los factores que influyen en la calidad de las actividades de evaluación, la ausencia de conocimiento matemático, interviene directamente en este caso (Cáceres y Chamoso, 2015).

Estos resultados son producto del estudio del proceso de evaluación seguido por una sola de las ocho parejas de futuros docentes involucradas en el estudio. El método de análisis, especialmente las categorías utilizadas, resultó útil para tratar de establecer tipos de comportamiento frente al proceso de evaluación entre pares. ¿Todos los estudiantes mostrarán ese apego a su propia resolución? ¿O dependerá esto de las diferencias que existan entre la propia solución y la solución evaluada? Estas preguntas guían los próximos pasos a seguir en esta investigación.

Cuanto más involucrados están los alumnos es su evaluación, más competentes son (Chanudet, 2016). Dado que la evaluación por pares involucra de forma directa a los estudiantes en su propio aprendizaje, puede contribuir a aumentar su motivación y su sentido de la responsabilidad (Lavy y Shriki, 2014). Además, en este caso concreto, se ha podido observar que la pareja evaluadora se encuentra con serias dificultades a la hora de comprender la resolución de su compañero. Por ello, las actividades de evaluación por pares pueden ayudarles no solo a desarrollar su pensamiento crítico y sus habilidades de análisis sino, en este sentido, sus conocimientos sobre el contenido trabajado y sus habilidades reflexivas (Beaver y Beaver, 2011), pues el hecho de que Carlos sea un par contribuye a que los futuros docentes reflexionen sobre los conocimientos y habilidades que debería tener en comparación con un igual. La evaluación por pares exige habilidades sociales y de comunicación, por lo que también puede contribuir a desarrollar sus capacidades sociales y de trabajo en equipo: aprender a dar y aceptar críticas, justificar sus opiniones y rechazar sugerencias (Lavy y Shriki, 2014). Por lo que, la evaluación entre pares, en particular con futuros docentes de Matemáticas, es una herramienta que tiene un gran potencial como herramienta de aprendizaje. Una de las preocupaciones expresadas por los docentes en formación es su falta de experiencia práctica con la evaluación, tanto en los cursos de nivel universitario como en las prácticas en colegios. La actividad que aquí se plantea puede ser un ejemplo a explotar en la formación docente.

Agradecimientos

Este trabajo ha sido subvencionado por el Proyecto de Investigación del Plan Nacional del MICINN con Referencia EDU2017-84276-R.

Referencias

- Arnal-Bailera, A., Muñoz-Escolano, J. M. y Oller-Marcén, A. M. (2018). Análisis de las anotaciones realizadas por profesores al calificar pruebas escritas de matemáticas. En L. J. Rodríguez-Muñiz, L. Muñiz-Rodríguez, A. Aguilar-González, P., Alonso, F. J. García García y A. Bruno (Eds.), *Investigación en Educación Matemática XXII* (pp. 131-140). Gijón: SEIEM.
- Beaver C. y Beaver. S. (2011). The Effect of Peer-Assessment on the Attitudes of Pre-Service Elementary and Middle School Teachers about Writing and Assessing Mathematics. *IUMPST: The Journal*, 5.
- Cáceres, M. J. y Chamoso, J. M. (2015). La evaluación sobre la resolución de problemas de matemáticas. En L. Blanco, J. A. Cárdenas y A. Caballero (Eds.), *La resolución de problemas de Matemáticas en la formación inicial de profesores de Primaria* (pp. 225-241). Cáceres: Universidad de Extremadura.
- Cárdenas, J. A., Blanco, L. J., Guerrero, E. y Caballero, A. (2016). Manifestaciones de los profesores de matemáticas sobre sus prácticas de evaluación de la resolución de problemas. *Bolema*, 30(55), 649-669.
- Chanudet, M. (2016). Assessing inquiry-based mathematics education with both a summative and formative purpose. En Peter Liljedahl y Manuel Santos-Trigo (Eds.) *Mathematical Problem Solving. Current Themes, Trends, and Research* (pp. 177-207). Suiza: Springer.
- Creswell, J. W. (2012). *Educational Research. Planning, conducting and evaluating quantitative and qualitative research*. Boston: Pearson.
- Custodia, E., Márquez, C. y Sanmartí, N. (2015). Aprender a justificar científicamente a partir del estudio del origen de los seres vivos. *Enseñanza de las ciencias*, 33(2), 133-155.
- Lavy, I. y Shriki, A. (2014). Engaging prospective teachers in peer assessment as both assessors and assessees: The case of geometrical proofs. *International Journal for Mathematics Teaching and Learning*.
- Nortvedt, G., Santos, L., y Pinto, J. (2016). Assessment for learning in primary school mathematics teaching: The case of Norway and Portugal. *Assessment in Education: Principles, Policy & Practice*, 23(3), 377-395.
- Schoenfeld, A. (1985). *Mathematical Problem Solving*. Orlando: Academic Press.
- Shahbari, J.A. y Abu-Alhija, F.N. (2018). Does Training in Alternative Assessment Matter? The Case of Prospective and Practicing Mathematics Teachers' Attitudes Toward Alternative Assessment and Their Beliefs About Nature of Mathematics. *International Journal of Science and Mathematics Education* 16(7), 1315-1335.
- Silver, E. A., y Mills, V. L. (Eds.) (2018). *A Fresh Look at Formative Assessment in Mathematics Teaching*. The National Council of Teachers of Mathematics.
- Suurtam, C., Koch, M. y Arden, A. (2010). Teachers' assessment practices in mathematics: classrooms in the context of reform. *Assessment in Education: Principles, Policy & Practices*, 17(4), 399-417.
- Topping, K.J. (2009). Peer assessment. *Theory into Practice*, 48(1), 20-27.
- Wyatt-Smith, C., Klenowski, V. y Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: a study of standards in moderation. *Assessment in Education: Principles, Policy & Practice*, 17(1), 59-75.