

Diseño y validación de un instrumento para observación de clases de matemáticas en Educación Secundaria: grupo nominal y método Delphi

Design and validation of an observation instrument for mathematics classes in secondary education: Expert panel and Delphi process

Arteaga-Martínez, B. ⁽¹⁾ , Macías-Sánchez, J. ⁽²⁾ , Pla-Castells, M. ⁽³⁾ ,
& Ramírez-García, M. ⁽⁴⁾ 

⁽¹⁾ Universidad Nacional de Educación a Distancia, UNED (España), ⁽²⁾ Universidad Complutense de Madrid (España),
⁽³⁾ Universitat de València (España), ⁽⁴⁾ Centro Universitario La Salle (España).

Resumen

El uso de los procesos de observación e interpretación de la práctica docente como herramienta de aprendizaje para los docentes es infrecuente en España. Para promover su utilización, se considera la importancia de contar con instrumentos que faciliten este análisis de la práctica, a la vez que recogen las características particulares del contexto educativo donde se utilizan. Este estudio presenta el proceso de diseño, construcción y validación de POEMat.ES, un instrumento de observación e interpretación de la práctica docente, en las clases de matemáticas de la etapa de Educación Secundaria. Este proceso consta de dos fases: 1) revisión de la literatura y diseño, utilizando un grupo nominal con 24 especialistas; 2) construcción y validación mediante un método Delphi en dos rondas, con 15 expertos. La versión final del instrumento se organiza en torno a tres dimensiones: contenido matemático, didáctica del contenido matemático y gestión del aula, con un total de 17 indicadores, cada uno de los cuáles se clasifica en cuatro niveles. Las metodologías de consenso utilizadas se consideran adecuadas para la elección y validación de los indicadores que enmarcan la observación de la práctica docente. Se concluye que POEMat.ES es una herramienta útil para sistematizar la observación de la práctica docente y así comprender e interpretar con precisión el proceso de enseñanza de las matemáticas en el aula de matemáticas de Secundaria según la realidad española.

Palabras clave: método Delphi, grupo nominal, enseñanza de las matemáticas, educación secundaria, instrumento de observación.

Abstract

Observation and interpretation processes are rarely used in teaching practice as learning tools in Spain. In order to encourage their use, it is important to have instruments that facilitate analysis in practice and that gather information about the particular characteristics of the educational context in which they are used. The present study presents the process of design, construction and validation of POEMat.ES, an observation and interpretation instrument for use in relation to the teaching of secondary education mathematics classes. This process consisted of two phases: 1) literature review and design using an expert panel made up of 24 specialists; 2) construction and validation via the Delphi method conducted by 15 experts over two rounds. The final version of the tool was organized around three dimensions: mathematical content, didactics of mathematical content and classroom management. It comprised 17 indicators, each of which was classified according to four levels. The methodology applied conformed to consensus, with this being considered adequate for the selection and validation of the indicators that frame the observation of teaching practice. It can be concluded that POEMat.ES is a useful tool for systematizing the observation of teaching practice in order to understand and accurately interpret the process of teaching mathematics in secondary education in the Spanish context.

Keywords: Delphi method, expert panel, mathematics teaching, secondary teaching, observation tool.

Received/Recibido

Jul 20, 2021

Approved /Aprobado Nov 29, 2021

Published/Publicado

Dec 22, 2021

La investigación en educación matemática ha encontrado un foco de análisis en la observación de aula que la ha convertido en una potente herramienta (Bostic et al., 2021) para el desarrollo profesional, dado el enorme impacto del desempeño docente en los niveles de rendimiento de los estudiantes (Weber et al., 2018).

Así nuestro trabajo se inicia haciéndonos una pregunta, ¿qué pautas de observación se utilizan en el aula de matemáticas en Educación Secundaria?

En una primera revisión de literatura especializada en didáctica de las matemáticas, se realizó una búsqueda de instrumentos de observación. El primer resultado concluyente de esta revisión fue la escasez de pautas específicas para la educación matemática, empleándose instrumentos generales de distintas materias (Schlesinger & Jentsch, 2016), situación que puede restar sensibilidad a la observación. Así, encontramos pautas de observación que recogían aspectos de la gestión de aula, la atención específica y la activación cognitiva, estos dos últimos referidos al estudiante. Estas tres dimensiones se han tenido en cuenta en distintas investigaciones, que al particularizar en la enseñanza de las matemáticas reflejan carencias en la recogida de información referida al dominio de la materia (Praetorius & Charalambous, 2018). Los instrumentos de observación que se analizaron en profundidad fueron: MQI *Mathematical Quality of Instruction* (Hill et al., 2008), IQA *Instructional Quality Assessment* (Matsumura et al., 2008) y la Pauta de observación de clases de matemáticas impartidas por profesores principiantes PROMATE (Barriendos et al., 2018).

Otros instrumentos analizados sirvieron también para hacer una primera aproximación a los focos de observación que podrían existir en un aula donde se enseñasen matemáticas: RTOP *Reform-Oriented Teaching Observation Protocol* (Piburn & Sawada, 2000), M-Scan *Mathematics Scan* (Walkowiak et al., 2014), *Elementary Mathematics Classroom Observation Form* (Thompson & Davis, 2014), e ISTOF *International System of Teacher*

Observation and Feedback (Teddlie et al., 2006).

Tras este primer análisis de herramientas de observación nos cuestionamos, ¿existen elementos propios de la cultura que influyen en los modelos de enseñanza de las matemáticas? La revisión mostró que existen pocas investigaciones (Thomas & Berry III, 2019) que aporten datos comparativos en contextos transculturales desde la perspectiva de la enseñanza, quizá debido al “sentido de la universalidad de los conceptos matemáticos” (Orton, 2003, p. 105). Sin embargo, desde el aprendizaje de las matemáticas, el análisis de las diferencias de carácter cultural se aprecia en distintos trabajos, por ejemplo, desde las creencias del estudiante (Diego-Mantecón & Córdoba-Gómez, 2019) o la resolución de problemas (Eccius-Wellmann et al., 2017). Por el contrario, con cierta frecuencia, la literatura científica de didáctica de las matemáticas muestra resultados que avalan las diferencias transnacionales referidas tanto a la didáctica como al aprendizaje (Openshaw & Walshaw, 2019).

Nos planteamos por tanto como objetivo general la construcción de una pauta de observación adaptada a la realidad de las aulas de secundaria en España, para facilitar el análisis de la práctica desde la perspectiva de la observación de las acciones del docente a través de la grabación de las sesiones de aula. Este artículo describe su construcción y validación, organizado en dos etapas, que consideramos como objetivos específicos:

1. Conformar un grupo de especialistas para la construcción de los primeros indicadores utilizando la técnica de grupo nominal (TGN).
2. Utilizar un método Delphi para el análisis y validación de los indicadores seleccionados que dé lugar a la pauta definitiva.

Método

Metodología en el diseño del instrumento

El estudio previo de la literatura, sobre instrumentos de observación en el aula de matemáticas, configuró una primera estructura en torno a tres dimensiones fundamentales, que cubren todo el espectro de análisis de la actuación del docente en el aula: el conocimiento matemático, el conocimiento didáctico de las matemáticas y la gestión de aula (Aguilar-González et al., 2018; Garzón, 2017; Van Zoest et al., 2021).

El equipo coordinador eligió un grupo de especialistas cuyo trabajo conjunto se iba a desarrollar de acuerdo a la TGN. El objetivo era construir los indicadores asociados a cada una de estas dimensiones.

Este grupo elegido desde una perspectiva transnacional y transcultural, trabajó de manera conjunta durante cuatro semanas, aprovechando una estancia colectiva en una universidad americana. El grupo estaba integrado por profesores de matemáticas de distintas universidades, teniendo en común que todos ellos habían participado como docentes en la formación de futuros docentes de matemáticas. Los docentes de este grupo procedían de distintos países: España (n=17), USA (n=3), Finlandia (n=3) e Israel (n=1). Se celebraron reuniones estructuradas cara a cara, para desde el conjunto de ideas individuales dar lugar a consensos en grupo.

El objetivo de esta etapa fue aumentar la productividad creativa del grupo para dar lugar a una decisión conjunta desde la estimulación de ideas críticas, a través de reuniones estructuradas que facilitasen a los participantes la oportunidad de que su postura se escuche (McMillan et al., 2016) a la vez que conocen las opiniones de otros miembros; una forma de involucrar a los propios actores en un fenómeno social para dar solución a un problema complejo de una manera colectiva (Moore, 1987). Las cuatro fases del TGN partieron de una estructura configurada previamente y comunicada a cada uno de los asistentes.

La primera fase de elaboración de respuestas individuales (generación silenciosa) partió de

preguntas sobre las tres dimensiones, de una en una y por escrito, para recoger la sensibilidad teórica de cada uno de los participantes. La segunda fase (*round robin*) sirvió para que el equipo coordinador solicitase la intervención de cada uno de los miembros cara a cara y así se generasen nuevas ideas. El equipo coordinador, mediante la videograbación de las sesiones, recopiló las ideas que luego presentaba resumidas al inicio de cada sesión, utilizando presentaciones digitales. La tercera fase centrada en la clarificación buscó la agrupación de ideas similares. Para ello, se utilizaban los apartados contruidos del instrumento para la observación y análisis de fragmentos de vídeo, grabados en distintas aulas de educación secundaria en España. Estas sesiones facilitaron que apareciesen cuestionamientos y conjeturas para el cambio de los indicadores. La última fase (votación) recogió de manera cuantitativa la valoración del grupo mediante el uso del instrumento de los fragmentos de vídeo; fue fundamental que esta fase se desarrollase de manera individual, digital y anónima.

Como resultado de este proceso, el instrumento de observación conservó la estructura inicial de 3 dimensiones; las dos primeras con 7 indicadores cada una y la tercera con 4. La puntuación numérica de 0 a 3 en cada uno de los indicadores no se ajusta a una escala ordinal, sino a un elemento clasificador en algunos de ellos.

La dimensión 1 “Contenido Matemático” pretende capturar las formas en las que el profesor articula el contenido intrínseco a la matemática a través del uso de las representaciones, la flexibilidad matemática, las definiciones y el razonamiento. Los tres primeros indicadores se fundamentan en los trabajos de Duval (1993) desde el uso, tratamiento (transformaciones del contenido matemático dentro del mismo registro) y conversión de representaciones. El cuarto valora la profundidad con la que el profesor define y usa las definiciones en el aula. El quinto se centra en los procesos de razonamiento que utiliza el profesor. El sexto, flexibilidad matemática, pretende observar si el profesor utiliza y promueve la capacidad de generar varias estrategias diferentes para una

tarea, idealmente comparándolas explícitamente y reflexionando sobre sus características (Schneider et al., 2011; Liu et al., 2018). Por último, un indicador recoge si el docente ha cometido error y, en caso afirmativo, su descripción.

La dimensión 2 “Didáctica del Contenido Matemático” busca estudiar los aspectos específicos de los procesos de enseñanza-aprendizaje de las matemáticas. “Uso de materiales” pretende observar si el profesor utiliza algún objeto, material o medio técnico que ayude al estudiante a comprender un contenido matemático, así como activar su práctica matemática (Maz-Machado et al., 2019). La naturaleza de las tareas propuestas identifica el tipo de tarea utilizada por el profesor dentro de una doble dicotomía determinada por el grado de indeterminación entre lo que se da y lo que se pide (Molina & Samper, 2019) y la accesibilidad, dependiendo si los estudiantes poseen el conocimiento matemático necesario para resolverlo o deben crearlo. Secuenciación y conexiones se centra en la cadena de tareas y su conexión con otros contenidos matemáticos en las argumentaciones realizadas por el profesor. La contextualización del contenido matemático pretende observar el uso de contextos “extramatemáticos” que dotan de significado a las nociones matemáticas y ayudan al estudiante a adquirir el campo conceptual sobre un contenido matemático, incluyendo el conjunto de situaciones y conceptos que lo conforman, mejorando su comprensión y funcionalidad (Kaiser, 2020; Vergnaud, 2013). Cesión de la responsabilidad de la actividad matemática busca capturar el grado de autonomía del estudiante, promovido por el profesor en el descubrimiento y validación de estrategias, fundamentado en la Teoría de Situaciones Didácticas (Brousseau, 2007). Lenguaje matemático recoge si el profesor usa un lenguaje formal que puede combinar palabras, símbolos y figuras con un significado matemático, es decir, determinado por unas reglas y relaciones muy precisas basadas en la deducción lógica, además de ajustarse al nivel educativo (Planas et al., 2018). La explotación didáctico-matemática de las intervenciones de

los estudiantes describe el uso que el profesor hace de las intervenciones de contenido matemático de los estudiantes. Estas situaciones son difíciles de prever (Rowland et al., 2009) y pueden aprovecharse para observar la calidad de la instrucción en el aula al considerar las oportunidades de los estudiantes para participar en debates matemáticos (Boston & Candela, 2018).

La dimensión 3 “Gestión del Aula” pretende observar aspectos de la práctica en el aula que intervienen en el proceso de enseñanza-aprendizaje, sin hacer referencia a aspectos específicos de la disciplina matemática. La densidad mide la razón del tiempo dedicado a la enseñanza-aprendizaje de las matemáticas respecto al tiempo total, sin valorar la calidad. Los trabajos previos de MQI muestran que la cantidad de tiempo que se dedica a la actividad matemática y a la gestión del aula juega un papel importante en la calidad de la enseñanza (Hill et al., 2008). Gestión de recursos pretende describir la claridad, eficiencia y eficacia del uso de recursos materiales. Uso del material escrito recoge si el profesor usa materiales escritos, hecho que en España parece cotidiano, y que en otros países no siempre existe esa dependencia del profesor hacia los libros de texto (van den Ham & Heinze, 2018). Por último, gestión de la conducta disruptiva pretende observar la efectividad de la gestión del profesor en una situación provocada por los estudiantes, que no favorece el desarrollo de una sesión (Rodríguez & Ruiz, 2019).

La validación del instrumento: el método Delphi

En la segunda fase de trabajo orientada a la validación de contenido se optó por el método Delphi (Martino, 1999). El objetivo de esta fase era asegurarse la inclusión de todos los elementos explicativos de la situación que se va a observar en un momento, donde la recogida de la información crítica puede resultar indispensable (Okoli & Pawlowski, 2004). Esta metodología facilita un pensamiento grupal que no se ve influido por opiniones dominantes de alguno de los miembros del panel.

Los instrumentos de recogida de información son cuestionarios de valoración de dimensiones e indicadores, intercalados con comentarios de opinión controlados. Se plantea, además, con un tiempo flexible, aunque limitado, que facilita la participación de un grupo disperso geográficamente y con distintos horarios (López-Gómez, 2018).

Selección y conformación del panel de expertos

El trabajo se inició con la formación del grupo de expertos, priorizando la representatividad de los participantes al número de los mismos. El muestreo que se utilizó para esta selección fue por criterio (Patton, 1990). Así, la selección del panel de expertos se hizo después de definir los criterios que dan respuesta a un equilibrio entre los conocimientos y percepciones actuales de la enseñanza de las matemáticas. Las dimensiones iniciales del instrumento establecieron como criterio fundamental la necesidad de que los expertos respondiesen a su pertenencia, de manera simultánea, a dos grupos: especialistas en matemáticas y formadores de docentes.

Con la necesidad de cubrir estos estratos con especialistas en la didáctica de las matemáticas se invitó a un total de 20 expertos, de distintas universidades españolas. En esa invitación se plantearon tres preguntas para valorar su capacidad predictiva, así como su compromiso con el proceso: años de experiencia profesional, descripción de su labor docente e investigadora, y una autoevaluación desde el objetivo del proyecto.

Para ello, se realizó el cálculo del índice de competencia experta (K) conformado por dos coeficientes: el índice de conocimiento (Kc) y el índice de argumentación (Ka). Este índice permite discriminar adecuadamente la inclusión o no de los expertos seleccionados en el panel (Llorente, 2013). Se consideró como criterio para la selección que K fuese superior a .8 (Cabero & Barroso, 2013).

Para obtener este valor, el equipo coordinador-investigador, realizó a criterio una ponderación de distintas fuentes de argumentación. Los pesos asignados sirven para valorar numéricamente el autoinforme que los expertos han completado. Para calcular Kc se solicitó a cada participante que valorase de 0 a 10 la siguiente cuestión: En una escala creciente de 0 a 10, ¿qué valor otorga usted a su nivel de capacitación (a través del conocimiento adquirido tanto desde su formación como su experiencia profesional) para abordar el proyecto de investigación planteado? Kc será esta puntuación multiplicada por .1. Para calcular Ka se recogieron las valoraciones de cada participante a la cuestión siguiente: ¿En qué grado (avanzado, medio o bajo) cada uno de los siguientes aspectos relacionados con su bagaje profesional serán importantes en sus argumentaciones en el desarrollo del proyecto de investigación? Estos aspectos se congregaron en una matriz de puntuación, organizada de acuerdo con la estructura prevista a priori para el instrumento de observación (Tabla 1), que da lugar a un resultado medio tras asignar la elección con el peso correspondiente.

Tabla 1. Matriz de puntuaciones para el cálculo del índice de argumentación

Temática	Avanzado	Medio	Bajo
1. Contenido matemático (sobre los temas que se abordan en la etapa considerada)	0.35	0.28	0.13
2. Contenido didáctico-matemático (sobre la enseñanza y el aprendizaje de las matemáticas)	0.30	0.24	0.12
3. Contenido pedagógico general	0.05	0.04	0.025
4. Formación inicial o permanente de profesores de matemáticas	0.10	0.08	0.075
5. Observación descriptiva de situaciones de aula (propias o de otros profesores)	0.10	0.08	0.075
6. Observación interpretativa de situaciones de aula (propias o de otros profesores)	0.10	0.08	0.075
Suma	1	0.80	0.50

No hay normas claras sobre el número óptimo para configurar estos paneles. La investigación previa señala que debe haber un número mínimo de 10 expertos (Parenté & Anderson, 1987). Landeta (1999) propone entre 7 y 30, y en la revisión hecha por Ludwig (1997) se indica que la mayoría de estudios han utilizado entre 15 y 20. Así, el grupo de expertos queda configurado por 15 personas, número considerado adecuado (Clayton, 1997). Se tuvo en cuenta el compromiso para mantenerse durante todo el proceso de validación (Buck et al., 1993).

El grupo de expertos son formadores de docentes en 7 universidades españolas, está integrado por 6 hombres y 9 mujeres, siendo un 86.67% doctores y un 93.33% licenciados en matemáticas.

Los expertos muestran una amplia heterogeneidad. Todos cuentan con experiencia docente en la universidad (M=17.86 años) y experiencia investigadora en el área de didáctica de las matemáticas (M=9.36 años). Además, siete de ellos cuentan con experiencia docente en la etapa de educación secundaria.

Las rondas del método Delphi

Los miembros del equipo coordinador son los encargados de recoger la información individual y anónima de cada uno de los expertos a través de un proceso iterativo, “devolviéndoles la propuesta del conjunto para su revisión y acuerdo” (Cabero & Llorente, 2013, p.17), así como elaboración de medidas de tendencia central de las valoraciones cuantitativas de los indicadores. La comunicación entre el equipo coordinador y cada uno de los panelistas se realiza por correo electrónico.

Este método asegura que cada uno de los expertos puede replantear la respuesta dada tras conocer las aportaciones de otros miembros del panel de manera que “este proceso iterativo avanza a la vez que se obtienen valores de consenso o acuerdo en los enunciados planteados o, si ello no se produce, se van consolidando posiciones en las que se observan discrepancias manifiestas” (López-Gómez, 2018, p. 26).

El número de rondas planificadas es de dos, por lo que puede considerarse como un método “Delphi modificado”, se busca así llegar a un consenso óptimo al tiempo que no se provoque fatiga en los participantes (Cabero, 2014), partiendo además de que los expertos trabajan sobre un tema que no han generado y sobre el que hay que llegar a acuerdos (Cabero & Infante, 2014). Dos rondas pueden considerarse un número adecuado para garantizar la convergencia (Linstone & Turoff, 1975).

Cabe valorar tanto el consenso como la discrepancia en este tipo de proceso respecto a cada uno de los elementos del instrumento, dado que cada aportación de los expertos se considerará importante (Gordon, 1994). Los juicios de los expertos se enriquecerán del proceso y la información que reciben, por lo que cabe pensar que las aportaciones últimas serán más ricas que las primeras, haciendo que la respuesta del grupo sea óptima (López-Gómez, 2018). Los expertos separados geográficamente, trabajaron de manera individual y anónima. La primera ronda se inició con el uso individual del instrumento sobre una muestra de fragmentos de vídeo de sesiones en aulas de Educación secundaria (cada investigador visualiza 29 fragmentos de entre 10-15 minutos) seleccionados por el equipo coordinador. Los expertos facilitaron información cualitativa y cuantitativa para cada uno de los indicadores del instrumento. El equipo coordinador envió un documento a cada uno de los miembros del panel, donde se expusieron los fragmentos de vídeo que debían visualizar. Estos fragmentos se seleccionaron de manera que cada experto visualizó fragmentos con las mismas características en cuanto a temas, curso, momento de la clase (inicio, medio o final) y docente. Este escenario de uso de la herramienta se considera el estímulo necesario para que los miembros del panel generen ideas y dé inicio el proceso de toma de decisiones. En esta primera ronda se pidió a cada uno de los expertos un aporte, desde sus revisiones de la literatura existente y el uso del instrumento. El cuestionario de evaluación se envió a cada uno de los expertos y se fijó un tiempo máximo para su

cumplimentación (de tres semanas). Este cuestionario solicitó a cada uno de los expertos una valoración en una escala de 0 a 10 de la pertinencia de cada uno de los indicadores del instrumento para responder a las preguntas de investigación planteadas. Se incluyó además una pregunta abierta, donde cada uno de los expertos podía incluir comentarios o justificar la puntuación de su valoración.

La segunda ronda parte del análisis de resultados de la primera ronda y el informe de modificaciones que cada uno de los expertos recibió. Este documento incluyó las modificaciones que se habían recogido mediante la clasificación de los aportes de la ronda 1 en condiciones de anonimato en las respuestas individuales. Los datos que se recogen parten de preguntas abiertas sobre la estructura del instrumento, pretendiendo en este sentido garantizar que el equipo coordinador no haya omitido algún tema relevante. “Al recibir la segunda ronda, se pide a los expertos considerar la posición de la medida de tendencia central del panel y se les permite revisar sus respuestas iniciales si así lo desean” (Novakowski & Wellar, 2008, p. 1486).

El tiempo entre rondas fue de tres meses.

Los resultados finales de esta ronda definieron el consenso para dar lugar al instrumento definitivo.

Resultados

Para la presentación de resultados finales del panel de expertos, se optó por una presentación gráfica de los resultados numéricos de las dos rondas, así como la presentación textual del instrumento, con un documento explicativo desde el diseño a los aportes en cada una de las rondas Delphi.

Los resultados de la primera ronda (Tabla 3) muestran valores similares en las dos primeras dimensiones y ligeramente inferiores en la tercera dimensión que, además, muestra una mayor dispersión en sus indicadores. Dos dimensiones (<15%) tienen una dispersión baja, mientras que una de ellas (15-30%) tiene una dispersión media (Martins & Theóphilo, 2007).

El análisis de los datos de la ronda 1 dio lugar a la petición a algunos de los expertos para que justificasen puntuaciones bajas en algunos de los indicadores, cuando la dispersión del valor era alta respecto a la opinión mayoritaria del grupo (Clayton, 1997). Los resultados de esta ronda facilitaron que el equipo de coordinación construyese una nueva versión del instrumento. Además de los cambios para cada indicador, se recogieron cambios de carácter estructural como la necesidad de definición de los términos tema y tarea, que se incluyeron en la introducción del instrumento. Todo el proceso detallando la intervención por indicador y de acuerdo a la ronda donde tuvo lugar se recoge en la Tabla 2. Cabe mencionar de manera expresa la supresión de un indicador tras la primera ronda.

Se facilitó a cada uno de los expertos la media muestral del grupo y una representación gráfica con la distribución de frecuencias y la respuesta individual del experto en la primera ronda, además de un desglose detallado de lo sucedido en cada una de las dimensiones de manera global.

Tabla 2. Resumen de cambios en el instrumento en las dos rondas del método Delphi

Indicador final del instrumento	Cambios primera ronda	Cambios segunda ronda
Uso de representaciones (1.1)	Se describen los niveles aclarando los registros a tener en cuenta en cada uno de ellos, diferenciando el registro de lenguaje natural (LN). Se modifican los ejemplos para ayudar a la identificación del nivel de puntuación.	Se especifica que los distintos registros no deben ser para un mismo contenido.
Tratamiento de representaciones	Surgen dudas sobre la necesidad de mantenerla por la información que puede aportar y el grado de especialización que debe tener el evaluador sobre los registros de representación para aplicar el indicador. Se suprime.	
Conversiones de representaciones (1.2)	Se simplifica la descripción de los niveles en función del número de registros entre los que se realizan conversiones para un mismo contenido, indicando el sentido de la conversión y la relación con LN. Se mejora el texto de la aclaración breve y se modifica los ejemplos.	Se excluye la representación LN.
Definiciones (1.3)	Se reduce al proceso de definir del profesor. Se modifica la aclaración breve para concretar el objetivo del indicador. Se modifican los niveles de puntuación y los ejemplos.	Se valora el rigor con el que un profesor define, cumpliendo las características propias de una definición matemática. Se modifica la aclaración breve, descripción de los niveles y ejemplos, facilitando aspectos como que el uso de una definición en una argumentación se considera en 1.4.
Argumentación (1.4)	Se cambia el nombre de razonamiento a argumentación. Se aclara el objetivo del indicador y se explica en la aclaración breve. Se modifican las descripciones de los niveles y ejemplos.	Se incluye el caso del uso de definiciones en la aclaración breve.
Flexibilidad matemática (1.5)	Se modifica la aclaración breve y los niveles de puntuación para clarificar las dificultades que se señalan en la puntuación. Se modifica algún ejemplo para facilitar la puntuación.	
Conexiones (1.6)	El indicador estaba incluido en la dimensión 2, con el nombre “Secuenciación y conexiones”, pasa a llamarse únicamente “Conexiones” y se sitúa en la primera dimensión, recogiendo las conexiones dentro de los temas matemáticos. Se modifican la aclaración breve y los niveles, eliminando todo lo relacionado con secuenciación y redactando mejor lo referido a conexiones dentro de las Matemáticas. Nuevos ejemplos.	Se describen con más precisión los niveles para recoger solo las conexiones utilizadas por el profesor entre temas y fenómenos matemáticos. Se modifica la aclaración breve, niveles y ejemplos.
Errores matemáticos del profesor (1.7)	Se reformula de manera íntegra, pasando de una situación dicotómica de sí/no, a una categorización que recoge los errores matemáticos del profesor al dar respuesta a las intervenciones de los alumnos. Se muestran los distintos tipos de errores a observar. No hay niveles, solo ejemplos.	Se incluyen 6 niveles, donde 0 es ausencia de errores y 1-5 tipos de errores matemáticos que comete el profesor al dar respuesta a los alumnos.
Uso de materiales (2.1)	Se modifican los niveles para centrarse en las acciones que realiza el docente con el material.	Se realizan leves modificaciones en la aclaración breve para distinguir los materiales de uso de material escrito incluido en la dimensión 3.
Naturaleza de las tareas propuestas (2.2.)	Se aclaran las nociones de “accesible” o “no accesible”, “cerrada” y “abierta”.	Se incluye en el nombre la palabra “propuestas”.
Contextualización del contenido matemático (2.3)	Se modifica para observar únicamente la contextualización del contenido matemático fuera del área. No se observa secuenciación.	Se especifica en el nivel 0 que no debe aparecer contextualización del contenido en todo el fragmento. Los ejemplos se redactan para poner el foco en el profesor.

(Tabla 2. Continuación)

Indicador final del instrumento	Cambios primera ronda	Cambios segunda ronda
Cesión de la responsabilidad de la actividad matemática (2.4)	Se modifica la descripción de los niveles para identificar si hay cesión de la responsabilidad y en qué términos.	Se añaden ejemplos y se mejoran acorde a la nueva descripción de los niveles.
Adecuación del discurso (2.5)	Se modifica el nombre de “Lenguaje Matemático” a “Adecuación del discurso”. Se centra en la adecuación del lenguaje al nivel educativo. Se modifica la aclaración breve, con una descripción de niveles y ejemplos.	
Explotación de las intervenciones de los estudiantes (2.6)	Se amplía la aclaración breve y se modifican levemente los ejemplos para concretar aspectos sobre calidad, intensidad y otros aspectos comunes en la descripción de los niveles. Se modifica el nivel 0 para recoger la ausencia de intervenciones o intervenciones sin contenido matemático.	Se amplía la aclaración breve y se modifican los ejemplos para concretar aspectos sobre calidad e intensidad.
Densidad (3.1)	Se modifican los ejemplos para ayudar a distinguir las puntuaciones entre 1-3.	Se definen los niveles en intervalos de tiempo dedicado a contenido matemático sin entrar en intensidad y calidad, incluido el trabajo autónomo del alumnado. Se modifica aclaración breve, descripción de niveles y ejemplos.
Uso de recursos expositivos (3.2)	Se cambia el nombre de “gestión de recursos” a “uso de recursos expositivos”, para enfocarla a la claridad de la exposición.	Se modifica la descripción de los niveles y la aclaración breve para precisar el enfoque hacia la claridad e inteligibilidad de la exposición. Cambios en las imágenes de los ejemplos.
Uso del material escrito (3.3)		Se indica en la aclaración que se valora solo lo que se pueda observar el uso del material escrito, no intensidad. La descripción de los niveles se modifica para recoger toda la casuística con el material escrito. Se modifican de los ejemplos.
Gestión de la conducta disruptiva (3.4)		Se identifican conductas disruptivas que dificultan los procesos de enseñanza-aprendizaje en el aula. Se valora la efectividad en la resolución de conflictos del profesor. Se modifica la aclaración breve, texto de los niveles y los ejemplos para ajustarlos al nuevo enfoque.

Nota. La pauta completa se puede consultar en <https://roderic.uv.es/handle/10550/78572>

Tras efectuar las modificaciones pertinentes y continuar con el proceso, los resultados de la segunda ronda señalan una mejora en el consenso (Tabla 3). La valoración

media en cada dimensión, así como la global, se han incrementado y se ha reducido la dispersión.

Tabla 3. Descriptivos por dimensión en las dos rondas Delphi

Dimensión	Primera ronda			Segunda ronda		
	M	s	CV	M	s	CV
1. Contenido Matemático	9.181	1.017	11%	9.495	0.652	7%
2. Didáctica del Contenido Matemático	9.167	0.974	11%	9.556	0.689	7%
3. Gestión del Aula	8.400	1.330	16%	9.200	0.840	9%
Total	8.992	1.129		9.447	0.724	

El método Delphi implica realizar los cálculos necesarios para determinar la variación en las valoraciones de cada experto

entre las distintas rondas. Para ello, iniciamos la exposición por los datos descriptivos de los indicadores (Tabla 4).

Tabla 4. Estadísticos descriptivos por indicador en las dos rondas Delphi

Indicador	Primera ronda						Segunda ronda					
	M	s	CV	Mo	Mdn	RI	M	s	CV	Mo	Mdn	RI
1.1.	9.5	0.6	0.1	10	10	1.0	9.7	0.5	0.1	10	10	0.5
1.2.	9.3	0.7	0.1	10	9	1.0	9.5	0.5	0.1	10	10	1.0
1.3.	8.9	1.1	0.1	10	9	2.0	9.3	0.7	0.1	9	9	1.0
1.4.	9.1	0.9	0.1	10	9	2.0	9.6	0.5	0.1	10	10	1.0
1.5.	9.3	1.1	0.1	10	10	1.0	9.6	0.6	0.1	10	10	1.0
1.6.	8.6	1.4	0.2	10	9	2.0	9.1	0.8	0.1	10	9	1.5
1.7.	9.4	0.9	0.1	10	10	1.0	9.6	0.7	0.1	10	10	0.5
2.1.	8.9	1.4	0.2	10	9	2.0	9.4	0.8	0.1	10	10	1.0
2.2.	9.1	0.8	0.1	9	9	1.5	9.7	0.6	0.1	10	10	0.5
2.3.	9.3	0.9	0.1	10	9	1.0	9.4	0.8	0.1	10	10	1.0
2.4.	9.3	0.8	0.1	10	9	1.0	9.6	0.6	0.1	10	10	1.0
2.5.	9.1	1.2	0.1	10	9	1.5	9.4	0.7	0.1	10	10	1.0
2.6.	9.5	0.6	0.1	10	10	1.0	9.9	0.4	0.0	10	10	0.0
3.1.	8.8	1.0	0.1	9	9	1.5	9.3	0.9	0.1	10	10	1.0
3.2.	8.3	1.0	0.1	8	8	1.0	9.1	0.7	0.1	9	9	1.0
3.3.	7.9	2.0	0.3	8	8	1.5	8.9	1.1	0.1	9	9	2.0
3.4.	8.6	1.1	0.1	8	9	1.0	9.5	0.5	0.1	9	9	1.0

Observamos que el consenso en las valoraciones entre rondas converge. Los comentarios que podemos hacer referidos a estos estadísticos son:

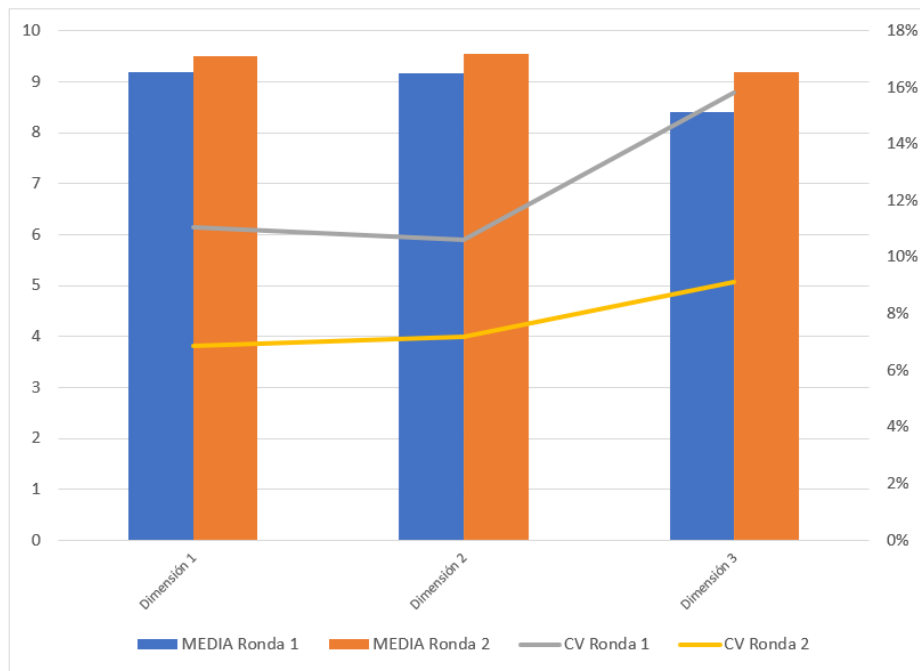
1. Rango intercuartílico (RI). El cálculo del RI es una de las herramientas que se utilizan en el método Delphi (von der Gracht, 2012). En nuestro caso, los RI de la segunda ronda son menores que en la primera y podemos considerar que se ha alcanzado un equilibrio, excepto el caso del indicador 3.3 que presenta un mayor RI en la segunda ronda. Para poder utilizar un nivel de aceptación único determinado de manera previa para todos los indicadores (Landeta, 1999) se ha utilizado el rango intercuartílico relativo (RIR), definido como el cociente entre RI y la mediana. De esta manera se consigue relativizar el RI por un valor de tendencia central. Se considera estabilidad de grupo si la variación en el RIR entre rondas es inferior a .30 (Mengual et al., 2016), situación que se cumple en todos los indicadores.

2. El coeficiente de variación (CV), una medida sin unidades, que se considera el mejor parámetro para este tipo de estudios (Shah & Kalaian, 2009), ha disminuido en todos los casos por lo que la dispersión ha convergido entre ambas rondas. El indicador 3.3 es de nuevo quien presenta mayor dispersión.

Estos resultados, acompañados por las medias y modas de cada uno de los indicadores, nos sitúan en unos resultados óptimos de consenso y estabilidad, lo que permite detener el proceso tras esta segunda ronda.

Los datos conjuntos (Figura 1) muestran una mejoría en los dos sentidos, un incremento de la valoración media de cada uno de los indicadores, al tiempo que una dispersión menor en las valoraciones. El CV es inferior al 15% en todos los casos, cumpliendo así que con un $CV > 30\%$ los datos deben ser descartados (Martínez, 1988).

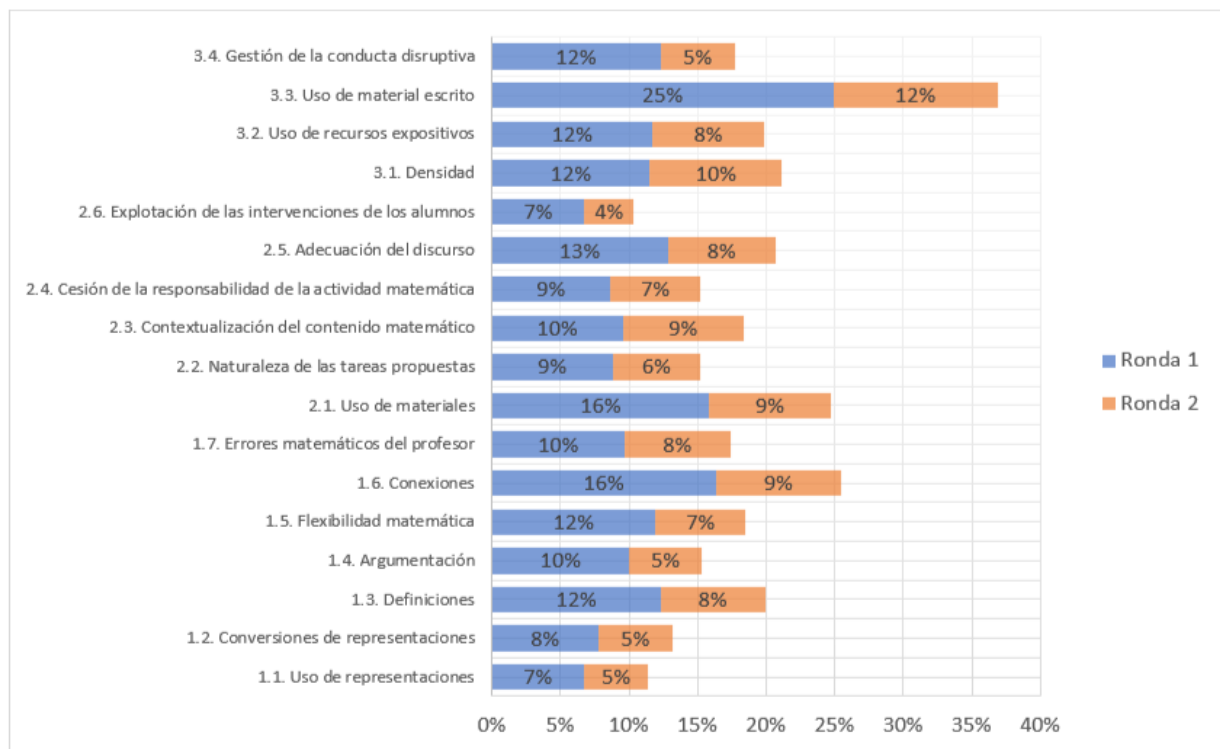
Figura 1. Media y coeficiente de variación de cada una de las rondas por dimensión



Para ver la información detallada en cada uno de los indicadores, dado que el objetivo es una búsqueda de consenso, recogemos cuál ha

sidido la variabilidad en cada uno ellos (Figura 2).

Figura 2. Dispersión en la valoración de los indicadores en las rondas 1 y 2



Conclusiones

Retomando las cuestiones iniciales que dieron lugar a los objetivos de este trabajo, señalamos que son pocas las pautas de observación para el aula de matemáticas en Educación Secundaria, aun siendo un objeto de investigación emergente al que se está dedicando esfuerzos en la comunidad investigadora, como pudimos ver en el número especial de la revista ZDM (Charalambous & Praetorius, 2018). El diseño de un instrumento que facilite la descripción de la práctica docente del profesor de matemáticas puede ser un elemento necesario para alcanzar el objetivo de mejorar los programas de formación, al tiempo que servir como herramienta de reflexión.

La construcción de una herramienta que guíe el análisis de la práctica docente a través de la grabación, análisis e interpretación de las sesiones de aula constituye una base necesaria para la formación inicial y la mejora docente desde el punto de vista de la didáctica y el análisis del contenido matemático.

Ante la segunda de las cuestiones, y tras haber trabajado con especialistas de distintos países que han facilitado la configuración de la estructura de POEMat.ES, podemos señalar que sí existen elementos de la cultura que influyen en los modelos de enseñanza de las matemáticas.

El enfoque conjunto desde la didáctica específica y las particularidades subyacentes al contenido de las matemáticas, contribuyen a un análisis especializado que pone el foco en el docente como vehículo que conecta los contenidos curriculares con las posibilidades de aprendizaje de sus estudiantes, además de situarnos en un marco particular como es el contexto español, que da lugar a fenómenos particulares como el uso del libro de texto como elemento regulador de la acción de enseñanza (Dolores & Ibáñez, 2020; Monterubio & Ortega, 2009). El primero de los objetivos específicos se ha logrado desde la revisión de instrumentos de observación de las clases de matemáticas, la configuración del grupo de especialistas y los resultados del

proceso llevado a cabo por el grupo nominal, que nos facilitaron configurar un instrumento adaptado al entorno escolar y social donde vamos a utilizarlo.

El instrumento construido se adapta a la realidad de las aulas en España y recoge los principales focos observados en la acción práctica desde tres dimensiones, contenido matemático, didáctica del contenido matemático y gestión del aula, cuya gestión depende fundamentalmente de la acción docente.

Los aspectos metodológicos en la construcción se han adaptado a las características del grupo de investigación y a la disposición de medios en cada momento, siendo coherentes con las pautas metodológicas requeridas para la consecución del objetivo, tanto en la primera parte de construcción con el grupo nominal, como en la segunda de mejora y validación, con el método Delphi.

La construcción del instrumento mediante el grupo nominal dio lugar a 17 indicadores distribuidos en tres dimensiones. Cada indicador se evalúa con una puntuación categórica de 0 a 3. Estos indicadores, durante el método Delphi se mantienen en número pero, sin embargo, se sitúan en dimensiones distintas o cambian su presentación, fundamentalmente desde la redacción o los ejemplos que ilustran cada uno de ellos, que ayudarán al docente a situar sus observaciones en una u otra categoría.

El segundo de los objetivos se ha logrado a raíz de la utilización del método Delphi, que ha dado lugar a un instrumento con evidencia de validez sólida, gracias a la cual “los investigadores en educación matemática pueden ser más capaces de analizar empíricamente las observaciones en el aula” (Bostic et al., 2021, p.9).

Las limitaciones planteadas en este trabajo provienen del número de expertos en el panel Delphi que, si bien desde la valoración global de proceso la consideramos suficiente, hubiese sido más rico contar con un mayor número de participantes. Sin embargo, cabe señalar como

positivo que el panel de expertos se ha mantenido activo desde su inicio, lo que da solidez y consistencia al trabajo realizado. La segunda limitación proviene del material en vídeo que se ha utilizado en la primera de las rondas Delphi para la valoración posterior del instrumento. Los expertos visualizaron los mismos fragmentos, elegidos de manera minuciosa para evitar los sesgos, pero es complejo en una investigación con personas que dichos fragmentos no hayan influido de ninguna manera en las posibles respuestas.

Esta investigación señala como prospectiva el uso de la pauta POEMat.ES en distintos cursos de Educación Secundaria en España, de manera que se puedan contrastar evidencias en los resultados, así como la construcción de un catálogo de buenas prácticas en el aula de matemáticas.

Financiación

Real Colegio Complutense (RCC) de Harvard

Agradecimientos

Queremos reconocer al Real Colegio Complutense (RCC) de Harvard por el apoyo institucional y financiero brindado para esta investigación como parte del Grupo de Estudio del RCC SIMIS: <https://rcc.harvard.edu/studying-and-improving-mathematics-instruction-secondary-schools-spain-simis>

Referencias

Aguilar-González, A., Muñoz-Catalán, M.C., Carrillo-Yáñez, J., & Rodríguez-Muñiz, L. J. (2018). ¿Cómo establecer relaciones entre conocimiento especializado y concepciones del profesorado de matemáticas? *PNA*, 13(1), 41-61. <http://dx.doi.org/10.30827/pna.v13i1.7944>

Barriandos, A., Berger, B., Domínguez, E., & Martínez, M. V. (2018). *Manual PROMATE. Pauta de observación de clases de matemáticas impartidas por profesores principiantes*. Centro de Investigación Avanzada en Educación de Chile.

Bostic, J., Lesseig, K., Sherman, M., & Boston, M. (2021). Classroom observation and mathematics education research. *Journal of Mathematics Teacher Education*, 24, 5–31. <https://doi.org/10.1007/s10857-019-09445-0>

Boston, M.D., & Candela, A.G. (2018). The Instructional Quality Assessment as a tool for reflecting on instructional practice. *ZDM Mathematics Education*, 50(3), 427-444. <https://doi.org/10.1007/s11858-018-0916-6>

Brousseau, G. (2007). *Iniciación al estudio de la teoría de las situaciones didácticas*. Libros del Zorzal.

Buck, A. J., Gross, M., Hakim, S., & Weinblatt, J. (1993). Using the Delphi process to analyze social policy implementation: A post hoc case from vocational rehabilitation. *Policy Sciences*, 26, 271-288. <https://doi.org/10.1007/BF00999473>

Cabero, J. (2014). Formación del profesorado universitario en TIC. Aplicación del método Delphi para la selección de los contenidos formativos. *Educación XXI*, 17(1), 111-132. <https://doi.org/10.5944/educxx1.17.1.10707>

Cabero, J., & Barroso, J. (2013). La utilización del juicio de experto para la evaluación de TIC: el coeficiente de competencia experta. *Bordón. Revista de Pedagogía*, 65(2), 25-38. <https://doi.org/10.13042/brp.2013.65202>

Cabero, J., & Infante, A. (2014). Empleo del método Delphi y su empleo en la investigación en comunicación y educación. *EDUTEC. Revista electrónica de tecnología educativa*, 48, a272-a272. <https://doi.org/10.21556/edutec.2014.48.187>

Cabero, J., & Llorente, M. C. (2013). La aplicación del juicio de experto como técnica de evaluación de las tecnologías de la información y comunicación (TIC). *Revista de Tecnología de Información y Comunicación en Educación*, 7(2), 11-22.

Charalambous, C.Y., & Praetorius, A.K. (2018). Studying mathematics instruction through different lenses: setting the ground

- for understanding instructional quality more comprehensively. *ZDM Mathematics Education*, 50(3), 355–366. <https://doi.org/10.1007/s11858-018-0914-8>
- Clayton, M. J. (1997). Delphi: a technique to harness expert opinion for critical decision-making tasks in education. *Educational psychology*, 17(4), 373–386. <https://doi.org/10.1080/0144341970170401>
- Diego-Mantecón, J. M., & Córdoba-Gómez, F. J. (2019). Adaptación y validación del MRBQ (Mathematics-Related Beliefs Questionnaire) al contexto colombiano con estudiantes de secundaria. *Educación matemática*, 31(1), 66–91. <https://doi.org/10.24844/em3101.03>
- Dolores, C., & Ibáñez, G. (2020). Conceptualizaciones de la Pendiente en Libros de Texto de Matemáticas. *Bolema: Boletim de Educação Matemática*, 34, 825–846. <https://doi.org/10.1590/1980-4415v34n67a22>
- Duval, R. (1993). Registros de representación semiótica y funcionamiento cognitivo del pensamiento. En E. Hitt (Ed.), *Investigaciones en Matemática Educativa II* (pp. 173–201). Grupo Editorial Iberoamérica.
- Eccius-Wellmann, C., Lara-Barragán, A. G., Martschink, B., & Freitag, S. (2017). Comparación de perfiles de ansiedad matemática entre estudiantes mexicanos y estudiantes alemanes. *Revista iberoamericana de educación superior*, 8(23), 69–83. <https://doi.org/10.22201/iisue.20072872e.2017.23.246>
- Garzón, D. (2017). Análisis de las decisiones del profesor de matemáticas en su gestión de aula. *Revista Educación Matemática*, 29(3), 131–160. <https://doi.org/10.24844/EM2903.05>
- Gordon, T. J. (1994). The Delphi method. In J.C. Glenn & T.J. Gordon (Eds.). *Futures Research Methodology Version 2.0*. (CD-Rom). American Council for the United Nations University.
- Hill, H., Blunk, M.L., Charalambous, C.Y., Lewis, J.M., Phelps, G.C., Sleep, L., & Ball, D. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An Exploratory Study. *Routledge Cognition and Instruction*, 26(4). <https://doi.org/10.1080/07370000802177235>
- Kaiser, G. (2020) Mathematical Modelling and Applications in Education. In S. Lerman (Eds.), *Encyclopedia of Mathematics Education*. Springer. https://doi.org/10.1007/978-3-030-15789-0_101
- Landeta, J. (1999). *El método Delphi. Una técnica de previsión para la incertidumbre*. Ariel.
- Linstone, H. A., & Turoff, M. (1975). *The delphi method*. Addison-Wesley.
- Liu, R. D., Wang, J., Star, J. R., Zhen, R., Jiang, R. H., & Fu, X. C. (2018). Turning potential flexibility into flexible performance: Moderating effect of self-efficacy and use of flexible cognition. *Frontiers in psychology*, 9, 646. <https://doi.org/10.3389/fpsyg.2018.00646>
- Llorente, L. (2013). Assessing Personal Learning Environments (PLEs). An expert evaluation. *Journal of New Approaches in Educational Research*, 2(1), 39–44. <https://doi.org/10.7821/naer.2.1.39-44>
- López-Gómez, E. (2018). El método Delphi en la investigación actual en educación: una revisión teórica y metodológica. *Educación XXI*, 21(1), 17–40. <https://doi.org/10.5944/educXXI.20169>
- Ludwig, B. (1997). Predicting the future: Have you considered using the Delphi methodology. *Journal of extension*, 35(5), 1–4.
- Martínez, A. (1988). *Diseños experimentales: métodos y elementos de teoría*. Trillas.
- Martino, J. (1999). Thirty years of change and stability. *Technological Forecasting and Social Change*, 62(1–2), 13–18. [https://doi.org/10.1016/S0040-1625\(99\)00011-6](https://doi.org/10.1016/S0040-1625(99)00011-6)

- Martins, G. A., & Theóphilo, C. R. (2007). *Metodologia da Investigação Científica para Ciências Sociais Aplicadas*. Atlas
- Matsumura, L. C., Garnier, H., Slater, S. C., & Boston, M. (2008). Toward measuring instructional interactions 'at-scale'. *Educational Assessment*, 13(4), 267–300. <https://doi.org/10.1080/10627190802602541>
- Maz-Machado, A., Madrid, M. J., León-Mantero, C., & Jiménez-Fanjul, N. (2019). Mathematical practical sessions with manipulatives: Trainee teachers' perceptions of their utility. *South African Journal of Education*, 39(2), Art. #1620. <https://doi.org/10.15700/saje.v39ns2a1620>
- McMillan, S.S., King, M., & Tully, M.P. (2016). How to use the nominal group and Delphi techniques. *International Journal of Clinic Pharmacy*, 38, 655–662. <https://doi.org/10.1007/s11096-016-0257-x>
- Mengual, S., Roig, R., & Mira, J.B. (2016). Delphi study for the design and validation of a questionnaire about digital competences in higher education. *International Journal of Educational Technology in High Education*, 13, 12. <https://doi.org/10.1186/s41239-016-0009-y>
- Molina, O., & Samper, C. (2019). Tipos de problemas que provocan la generación de argumentos inductivos, abductivos y deductivos. *Bolema: Boletim de Educação Matemática*, 33, 109-134. <https://doi.org/10.1590/1980-4415v33n63a06>
- Moore, C. M. (1987). *Group techniques for idea building*. Sage Publications, Inc.
- Novakowski, N., & Wellar, B. (2008). Using the Delphi technique in normative planning research: methodological design considerations. *Environment and Planning A*, 40(6), 1485-1500. <https://doi.org/10.1068/a39267>
- Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & management*, 42(1), 15-29. <https://doi.org/10.1016/j.im.2003.11.002>
- Openshaw, R., & Walshaw, M. (2019). *Transnational Synergies in School Mathematics and Science Debates*. Springer Nature. <https://doi.org/10.1007/978-3-030-28269-1>
- Orton, A. (2003). *Didáctica de las matemáticas*. Ministerio de Educación, Cultura y Deporte-Ediciones Morata.
- Parenté, F. J., & Anderson, J.K. (1987). Delphi inquiry systems. En G. Wright & P. Ayton (Eds.), *Judgemental Forecasting* (pp. 129-156). John Wiley.
- Patton, M. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA.
- Piburn, M., & Sawada, D. (2000). *Reformed teaching observation protocol (RTOP): Reference manual. Technical Report*. Arizona Collaborative for Excellence in the Preparation of Teachers.
- Planas, N., Arnal-Bailera, A., & García-Honrado, I. (2018). El discurso matemático del profesor: ¿Cómo se produce en clase y cómo se puede investigar? *Enseñanza de las Ciencias*, 36(1), 45-60. <https://doi.org/10.5565/rev/ensciencias.2240>
- Praetorius, A. K., & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: looking back and looking forward. *ZDM Mathematics Education*, 50(3), 535-553. <https://doi.org/10.1007/s11858-018-0946-0>
- Rodríguez, J.M., & Ruiz, J. (2019). El clima social en centros educativos: percepción del profesorado de Educación Secundaria Obligatoria de la Comunidad de Madrid. *Revista de Investigación Educativa*, 37(1), 231-250. <https://doi.org/10.6018/rie.37.1.320541>
- Rowland, T., Turner, F., Thwaites, A., & Huckstep, P. (2009). Transformation: Using examples in mathematics teaching. In *Developing Primary Mathematics Teaching: Reflecting on Practice with the Knowledge Quartet* (pp. 67-100). SAGE

Publications.

<https://doi.org/10.4135/9781446279571>

- Schlesinger, L., & Jentsch, A. (2016). Theoretical and methodological challenges in measuring instructional quality in mathematics education using classroom observations. *ZDM Mathematics Education*, 48, 29–40. <https://doi.org/10.1007/s11858-016-0765-0>
- Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations between conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology*, 47(6), 1525–1538. <https://doi.org/10.1037/a0024997>
- Shah, H., & Kalaian, S. A. (2009). Which Parametric Statistical Method to Use For Analyzing Delphi Data? *Journal of Modern Applied Statistical Method*, 8(1), 226–232. <https://doi.org/10.22237/jmasm/1241137140>
- Teddlie, C., Creemers, B., Kyriakides, L., Muijs, D., & Yu, F. (2006). The international system for teacher observation and feedback: Evolution of an international study of teacher effectiveness constructs. *Educational Research and Evaluation*, 12(6), 561–582. <https://doi.org/10.1080/13803610600874067>
- Thomas, C. A., & Berry III, R. Q. (2019). A Qualitative Metasynthesis of Culturally Relevant Pedagogy & Culturally Responsive Teaching: Unpacking Mathematics Teaching Practices. *Journal of Mathematics Education at Teachers College*, 10(1), 21–30. <https://doi.org/10.7916/jmetc.v10i1.1668>
- Thompson, C. J., & Davis, S. B. (2014). Classroom observation data and instruction in primary mathematics education: Improving design and rigour. *Mathematics Education Research Journal*, 26(2), 301–323. <https://doi.org/10.1007/s13394-013-0099-y>
- van den Ham, A. K., & Heinze, A. (2018). Does the textbook matter? Longitudinal effects of textbook choice on primary school students' achievement in mathematics. *Studies in Educational Evaluation*, 59, 133–140. <https://doi.org/10.1016/j.stueduc.2018.07.005>
- Van Zoest, L. R., Peterson, B. E., Rougée, A. O., Stockero, S. L., Leatham, K. R., & Freeburn, B. (2021). Conceptualizing important facets of teacher responses to student mathematical thinking. *International Journal of Mathematical Education in Science and Technology*, 1–26. <https://doi.org/10.1080/0020739X.2021.1895341>
- Vergnaud, G. (2013). Pourquoi la théorie des champs conceptuels? *Infancia y Aprendizaje*, 36(2), 131–161. <https://doi.org/10.1174/021037013806196283>
- von der Gracht, H.A. (2012). Consensus measurement in Delphi studies: review and implications for future quality assurance. *Technological Forecasting and Social Change*, 79(8), 1525–1536. <https://doi.org/10.1016/j.techfore.2012.04.013>
- Walkowiak, T. A., Berry, R. Q., Meyer, J. P., Rimm-Kaufman, S. E., & Ottmar, E. R. (2014). Introducing an observational measure of standards-based mathematics teaching practices: Evidence of validity and score reliability. *Educational Studies in Mathematics*, 85(1), 109–128. <https://doi.org/10.1007/s10649-013-9499-x>
- Weber, K. E., Gold, B., Prilop, C. N., & Kleinknecht, M. (2018). Promoting pre-service teachers' professional vision of classroom management during practical school training: Effects of a structured online-and video-based self-reflection and feedback intervention. *Teaching and Teacher Education*, 76, 39–49. <https://doi.org/10.1016/j.tate.2018.08.008>

Authors / Autores

Arteaga-Martínez, Blanca (blanca.arteaga@edu.uned.es)  0000-0002-1079-1526

Licenciada en CC. Matemáticas por la Universidad Autónoma de Madrid y Doctora en CC. de la Educación por la Universidad Complutense de Madrid. Ha combinado su labor como profesora en niveles obligatorios con la docencia en la universidad (Universidad Carlos III de Madrid, Saint Louis University in Madrid, Universidad Internacional de La Rioja, Universidad de Alcalá y Universidad Rey Juan Carlos), además de desarrollar labores de coordinación de equipos. En la actualidad profesora Ayudante Doctora en la Universidad Nacional de Educación a Distancia (UNED). Su investigación se ha centrado en la investigación-acción en el aula, tomando la didáctica de la matemática como eje vertebrador. Ha dirigido y participado en distintos proyectos de investigación subvencionados y proyectos de innovación educativa con distintas universidades. Integrante del grupo de investigación Pedagogía Adaptativa de la Universidad Complutense de Madrid.

Macías-Sánchez, Jesús (j.macias@ucm.es)  0000-0001-9798-7654

Licenciado en CC. Matemáticas por la Universidad Complutense de Madrid, Máster en Formación del profesorado y Máster en estudios Avanzados en Pedagogía por la Universidad Complutense de Madrid y Doctor en CC. de la Educación por la Universidad Complutense de Madrid. Combina su labor como profesor en niveles obligatorios con la docencia en la universidad (Universidad Internacional de La Rioja, Universidad Europea de Madrid y Universidad Complutense de Madrid). En la actualidad profesor Asociado en la Universidad Complutense de Madrid (UCM) y profesor en Enseñanza Obligatoria en el Colegio Brotmadrid. Su investigación se ha centrado en la investigación-acción en el aula, fundamentado en el proceso de enseñanza-aprendizaje de las matemáticas. Ha participado en algún proyecto de investigación subvencionado y proyectos de innovación educativa.

Pla-Castells, Marta (marta.pla@uv.es)  0000-0002-2088-2536

Licenciada en CC. Matemáticas por la Universitat de València (UVEG) y Doctora en CC. de la Computación por la misma universidad. Ha trabajado como directora de proyectos europeos para el Instituto de Robótica y TIC de la UVEG mientras era profesora asociada en departamentos de la Facultad de Matemáticas de la UVEG y de la Universitat Jaume I de Castelló (UJI). En la actualidad es profesora contratada doctora en la UVEG. Por su experiencia en programación de software, ha sido capaz aplicar las nuevas tecnologías en el campo de la educación matemática. Ha trabajado también en formación continua de profesores de primaria en colaboración con el Centro de formación, innovación y recursos para el profesorado de la Generalitat Valenciana. En la actualidad pertenece al grupo de investigación de Modelización Matemática del departamento de Didáctica de la Matemática de la UVEG.

Ramírez-García, Mónica (mramirez@lasallecampus.es)  0000-0002-1198-2017

Licenciada en CC. Matemáticas por la Universidad Autónoma de Madrid y Doctora en CC. de la Educación por la Universidad Complutense de Madrid. Trabaja como profesora asociada en la Unidad de Didáctica de las Matemáticas de la Universidad Complutense de Madrid desde 2011 y como profesora titular en el Centro Universitario de La Salle desde 2006. Su investigación se ha centrado en la enseñanza-aprendizaje de las matemáticas en las primeras edades. Desde 2016 colabora con varios investigadores de la Universidad de Sevilla y de la Universidad de Huelva aplicando el modelo de conocimiento especializado del profesor de matemáticas (MTSK) a su actividad como formadora de maestros (especialmente de Educación Infantil en formación inicial y formación permanente) y como investigadora.



[ISSN: 1134-4032]

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



Esta obra tiene [licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/).

This work is under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).