

Análisis de los ítems de las evaluaciones autonómicas de diagnóstico en España: 2008-2009

Rosa Marta Caraballo Caraballo; Luis Rico Romero; José Luis Lupiáñez Gómez

Resumen

Este estudio se centra sobre las características, y su adecuación al modelo de competencias establecido por el estudio PISA de la OECD y por el Ministerio de Educación y Ciencia en España, de los ítems para la evaluación de diagnóstico en competencia matemática de los estudiantes de 2º de Educación Secundaria Obligatoria, elaborados por las Comunidades Autónomas españolas en el curso académico 2008-2009.

Abstract

This study focuses on the characteristics of the items for the diagnostics evaluation of mathematics competency of students in the 2nd course of the Compulsory Secondary Education, and its suitability for the competency model established by the OECD/PISA study and by the Ministry of Education and Science in Spain. These items were designed by the Spanish Autonomous Communities in the academic year 2008-2009.

Resumo

Este estudo centra-se nas características e sua adequação para o modelo de competências estabelecido pelo estudo PISA da OCDE e pelo Ministério da Educação e da Ciência da Espanha dos itens para a avaliação de diagnóstico em competência matemática dos alunos da 2º ano do ensino secundário obrigatório, desenvolvidas pelas comunidades autônomas espanholas no ano lectivo 2008-2009.

1. Introducción

Evaluar es una tarea necesaria e ineludible. Boulmetis y Dutwin (2000) definen la evaluación como “el proceso sistemático de recopilar y analizar datos con el propósito de determinar si, y en qué grado, se han logrado los objetivos propuestos” (p. 4). Un proceso evaluativo con un seguimiento adecuado provee datos para diagnosticar sobre los elementos participantes en una intervención particular y usa este diagnóstico para mejorar las actividades educativas subsiguientes. Evaluar es, de manera sucinta, recoger y sistematizar información para tomar decisiones. Las comunidades educativas que, conscientes de su importancia, han adoptado una cultura de evaluación poseen una herramienta poderosa para identificar y reflexionar respecto de las fortalezas y debilidades de sus prácticas educativas y para tomar decisiones que redunden en el mejoramiento del sistema y, por consiguiente, en el éxito de sus estudiantes.

En España, la Ley Orgánica 2/2006, del 3 de mayo, de Educación (LOE) establece en sus artículos 21 y 29 que las Administraciones Educativas de cada Comunidad Autónoma realizarán una evaluación de diagnóstico de carácter externo, censal y formativo a todos los alumnos que finalicen el segundo ciclo de la

Educación Primaria (9-10 años) y el segundo curso de la Educación Secundaria Obligatoria, ESO, (13-14 años), respectivamente (Ministerio de Educación y Ciencia, 2006). Se espera que estas evaluaciones contribuyan a mejorar la calidad y la equidad de la educación arrojando luz sobre la situación del sistema educativo mediante la valoración de los aprendizajes de los alumnos y el impulso de procesos de innovación y que informen, además, respecto del grado de adquisición y logro de las competencias básicas del currículo tanto a nivel nacional como regional (Artículo 140, LOE 2/2006).

En el estudio que aquí presentamos, analizamos los ítems de las primeras pruebas de diagnóstico en competencia matemática aplicados en España durante el año académico 2008-2009, tomando como referencia el modelo de las evaluaciones para matemáticas del *Programme for International Student Assessment* (PISA) de la *Organization for Economic Co-operation and Development* (OECD). Dicha organización estableció las evaluaciones PISA con el objetivo de determinar, entre otros, el grado de competencia matemática que los jóvenes de 15 años han adquirido para enfrentarse a situaciones de la vida cotidiana; uno de sus objetivos principales es establecer indicadores de calidad para determinar cómo los sistemas educativos logran ese nivel de formación (Rico, 2007). Las evaluaciones autonómicas de diagnóstico usan el modelo de evaluación en matemáticas establecido por el estudio PISA como marco de referencia para diseñar las pruebas (Lupiáñez, 2010).

Por el rol preponderante que desempeñan las competencias en la caracterización de las expectativas de aprendizaje, los resultados obtenidos por los estudiantes españoles en las pruebas PISA de 2003 y de 2006 incidieron en las nuevas orientaciones curriculares. Asimismo, tuvieron el efecto de enfatizar el carácter funcional del aprendizaje de las matemáticas escolares (Lupiáñez, 2009). Este enfoque apoya que los estudiantes en el nivel de Educación Obligatoria, pongan en juego el conocimiento teórico y el dominio técnico que han adquirido para resolver, de una manera eficaz y eficiente, una variedad de situaciones-problemas. Este enfoque ha sido consistentemente avalado por el proyecto PISA de la OECD. Tanto las recientes orientaciones curriculares en España a nivel nacional como el marco teórico de las evaluaciones de PISA, respaldadas ambas por el enfoque funcional del aprendizaje matemático, sirven de marco de referencia para esta investigación. Determinar el grado de ajuste al modelo PISA de los instrumentos elaborados por las comunidades autónomas españolas para atender a la evaluación de diagnóstico en segundo curso de la educación obligatoria, es el objetivo general que guía esta investigación (Caraballo, 2010).

2. La Ley Orgánica de Educación (LOE) y las evaluaciones de diagnóstico

La Ley Orgánica 2/2006, del 3 de mayo, de Educación (LOE) define el currículo como “el conjunto de objetivos, competencias básicas, contenidos, métodos pedagógicos y criterios de evaluación de cada una de las enseñanzas reguladas en la presente Ley” (Artículo 6). De esta manera se incorpora la noción de competencia como parte integral del currículo en todos los niveles educativos. Se considera la introducción del concepto de competencia como un cambio significativo en las normativas curriculares, ampliando así la orientación hacia la noción de objetivos del

aprendizaje que había imperado hasta el momento (Rico y Lupiáñez, 2008). Es importante destacar que las competencias y los objetivos difieren entre sí y entrañan expectativas distintas sobre el aprendizaje de las Matemáticas (Lupiáñez, 2009).

Sobre las ideas de aprendizaje permanente y de desarrollo de las competencias básicas la LOE introduce, como una novedad, la realización de dos procesos distintos para las evaluaciones de diagnóstico: las evaluaciones generales de diagnóstico y las evaluaciones autonómicas de diagnóstico. La evaluación autonómica de diagnóstico, cuya responsabilidad recae sobre las Administraciones Educativas, tendrá carácter formativo y orientador, ofrecerá información importante sobre la situación del alumnado, los centros y el sistema educativo como un conjunto y redundará en la adopción de planes de mejora de la calidad del sistema. Estas evaluaciones tendrán como marco de referencia las evaluaciones generales de diagnóstico creadas igualmente por virtud de la LOE. Ambos procesos evaluarán las competencias básicas del currículo, y generarán compromisos de revisión y mejora educativa con base en los resultados obtenidos (Artículo 144, LOE). Lo que intentan valorar las pruebas es el grado en que la escuela moldea al alumno para hacer frente a la vida y cómo lo prepara para desenvolverse como ciudadano, de una manera eficaz, en la sociedad a la cual pertenece. Es decir, que el conjunto de competencias básicas, según definidas en el currículo, se refiere a la capacidad del estudiante para aplicar los conocimientos, habilidades y actitudes adquiridos a la realidad y a la resolución de problemas enmarcados dentro del contexto de la vida cotidiana (Instituto de Evaluación, 2009).

3. El enfoque funcional del aprendizaje matemático

La evaluación PISA concibe las matemáticas como un conjunto de herramientas que sirven para resolver problemas mediante la puesta en funcionamiento de determinadas capacidades, conocimientos y estrategias (Rico, 2006). Tanto las nociones de alfabetización y competencia según se describe en el modelo PISA, como la percepción de las matemáticas como “modo de hacer”, se ajustan a un enfoque funcional del aprendizaje de las matemáticas (Rico y Lupiáñez, 2008). Este modelo centra su foco de atención en cómo los alumnos aplican los conocimientos adquiridos para enfrentarse a situaciones de la vida real que le son familiares. Son tres las componentes que conforman el modelo funcional: tareas contextualizadas, herramientas conceptuales y capacidades del sujeto cognitivo que trabaja. ¿Cómo interactúan estas tres componentes? El sujeto cognitivo usa las herramientas que tiene a su disposición para aproximarse a las tareas, manifestando su competencia al movilizar los correspondientes procesos cognitivos.

El modelo funcional requiere que el estudiante conozca las herramientas que tiene disponibles, pero su foco de atención principal reside en las tareas. La analogía que se establece es la siguiente: una vez el sujeto conoce las herramientas básicas de las matemáticas (conceptos y procedimientos), el énfasis se dirige, no a la herramienta misma como objeto de estudio, sino a su aplicación en distintas situaciones, a su empleo en satisfacer las demandas que plantean las tareas. Es decir, al uso de la herramienta en situaciones no convencionales. La orientación de funcionalidad dentro del currículo de Matemáticas no implica el conocimiento de los conceptos básicos, como herramientas, en todas sus manifestaciones sino la

extensión de su uso y aplicación a situaciones que así lo demanden (Rico y Lupiáñez, 2008).

4. El modelo matemático de PISA

El Programa PISA se considera un programa internacional completo y riguroso para evaluar el desempeño de los estudiantes y para recopilar información respecto de los factores relacionados con el estudiantado, las familias y las instituciones que permiten explicar las diferencias en tal desempeño (OECD, 2004). El dominio matemático que evalúan las pruebas PISA se conoce como alfabetización matemática (*mathematical literacy*) y, en términos generales, competencia matemática. Se describe este dominio como “las capacidades de los estudiantes para analizar, razonar y comunicar eficazmente cuando resuelven o enuncian problemas matemáticos en una variedad de dominios y situaciones” (Rico, 2006; p. 276). Esta definición gira en torno a usos más amplios de las matemáticas en la vida de las personas y no está limitada a las operaciones mecánicas o algorítmicas.

Inherente al concepto de competencia se halla la necesidad de que el estudiante sea capaz de adaptar y aplicar su comprensión matemática a una amplia gama de contextos dentro y fuera de la sala de clase. Lejos de enfatizar el dominio de conceptos y técnicas, la evaluación de PISA enfatiza la capacidad del estudiante para enfrentarse, mediante el uso de las matemáticas, a una amplia variedad de situaciones cotidianas. Por lo tanto, el foco de atención de PISA no está en lo que el estudiante ha aprendido mediante un currículo escolar específico y convencional sino en las destrezas y los conocimientos necesarios para actuar como un adulto informado, reflexivo y como miembro de una sociedad. La alfabetización matemática se evalúa aplicando a los estudiantes tareas basadas en situaciones que representan el tipo de problemas que enfrentarán en la vida real.

Dentro del marco conceptual del modelo matemático PISA se denomina matematización a la estrategia general que usan los estudiantes para resolver problemas de la vida real. Matematizar es un proceso fundamental e inherente a la alfabetización matemática (OECD, 2009). Cuando un estudiante se mueve entre el contexto real del problema y el mundo matemático necesario para resolverlo, está matematizando. Este proceso cíclico incluye la interpretación y la evaluación del problema así como la reflexión en su solución para verificar que la solución obtenida, en efecto, responde a la situación real que inicialmente lo originó.

4.1. Caracterización de las pruebas PISA

Las pruebas PISA se aplican a estudiantes que hayan cumplido 15 años sin tomar en consideración el nivel escolar que cursen en ese momento. En la caracterización de las evaluaciones matemáticas PISA se distinguen tres tipos de variables: variables de tarea, variables de sujeto y variables de resultado o de desempeño (OECD, 2004). Las variables de tarea delimitan los instrumentos de evaluación de las pruebas PISA para satisfacer el principio de matematización o competencia sobre el cual se fundamentan. Las variables de sujeto caracterizan a la muestra elegida y son las propias de la población española. Para llevar a cabo un análisis detallado de la competencia matemática se han identificado ocho competencias específicas –pensar y razonar, argumentar, comunicar, modelizar,

plantear y resolver problemas, representar, usar lenguaje técnico, simbólico y las operaciones y usar ayudas y herramientas– que son relevantes y significativas en todos los niveles escolares y que no responden a áreas o contenidos curriculares específicos. Estas competencias proveen la base para las escalas o niveles de rendimiento que constituyen las variables de desempeño en el proyecto PISA. A los efectos de nuestro estudio, centrado en los instrumentos de evaluación de las pruebas de diagnóstico, son las variables de tarea que delimitan el diseño de los ítems las que tienen mayor relieve. Las variables de sujeto deben atender a respetar las condiciones y requisitos generales del estudio PISA y a satisfacer los criterios de representatividad de las muestras. Las variables de resultado permitirán, en su caso, efectuar un control empírico de idoneidad de los instrumentos de evaluación.

4.2. Las variables de tarea en PISA como criterio para valorar los ítems

Las variables de tarea que consideran las evaluaciones PISA son: el contenido matemático al que se relacionan las tareas, las situaciones o contextos usados como estímulo para enmarcar el problema en el mundo real y la complejidad de las tareas que se proponen, agrupadas éstas según tres niveles de demanda cognitiva. Los problemas propuestos ocurren en una situación real; esta situación provee el contexto para la tarea matemática. A fin de usar las matemáticas para resolver el problema, el estudiante debe poseer un grado de dominio del contenido matemático relevante. Y a fin de resolver el problema debe desarrollarse y seguirse un proceso de resolución. Luego, para ejecutar este proceso exitosamente, el estudiante necesita determinadas competencias que el modelo agrupa al establecer tres niveles hipotéticos de complejidad en la demanda cognitiva que plantea la tarea. De esta manera el estudiante matemáticamente alfabetizado se enfrenta a un proceso de resolución de un problema enmarcado en un contexto, un contenido matemático y mediante una demanda que tiene un determinado nivel de complejidad. Vemos cómo el modelo PISA se ajusta al modelo funcional del conocimiento matemático conformado por unas tareas contextualizadas, unas herramientas conceptuales y el sujeto cognitivo que muestra sus capacidades (OECD, 2009).

En PISA el contenido no se considera como relativo a las áreas matemáticas tradicionales sino como ideas amplias o clave (*overarching ideas*) estructuradas alrededor de diferentes fenómenos que describen conceptos, ideas y estructuras matemáticas. No existe una correspondencia estricta entre estas ideas amplias y las áreas matemáticas tradicionales. Atemperadas a la orientación principal del currículo escolar, PISA considera las siguientes ideas amplias como contenidos matemáticos, a partir de categorías fenomenológicas:

- a. Cantidad. Este contenido responde a la Aritmética y abarca todo lo relativo a la cuantificación necesaria para organizar el mundo. En esta categoría se incluye el sentido numérico, la comprensión del significado de las operaciones, el sentido de la magnitud de las cantidades y los números, los cálculos mentales y la estimación.
- b. Espacio y forma. Este contenido responde a la Geometría tradicional e implica comprender, describir e interpretar el mundo físico que nos rodea. Se incluyen en esta categoría los siguientes aspectos: reconocer formas y modelos; describir, codificar y decodificar la información visual; comprender los cambios

dinámicos de las formas; establecer similitudes y diferencias; posiciones relativas; representaciones bidimensionales y tridimensionales y las relaciones entre ambas y la orientación en el espacio.

- c. Cambio y relaciones. Este contenido responde al Cálculo y al Álgebra y cubre las manifestaciones y los procesos de cambio en el mundo que nos rodea acometidos mediante el estudio de las relaciones que, a su vez, se abordan desde el ámbito de las funciones matemáticas. Para representar las relaciones pueden usarse los símbolos, las tablas y las gráficas y los dibujos geométricos.
- d. Incertidumbre. Este contenido responde a la Estadística y la Probabilidad y abarca el tratamiento de los datos y el azar. Las ideas principales que se incluyen son la recopilación, análisis y representación de los datos, la probabilidad y la inferencia.

Frecuentemente subestimado o ignorado en las matemáticas escolares, el contexto en el cual se sitúa un problema juega un papel importante en la resolución de problemas de la vida real y en la alfabetización matemática (OECD, 2009). PISA destaca la importancia del contexto y le otorga un papel principal en la evaluación de la alfabetización matemática. Puesto que el ciudadano común se relaciona con contextos de diversa naturaleza, PISA reconoce la necesidad de incluir en sus evaluaciones una amplia gama de contextos. Según su grado de cercanía con la situación particular del estudiante, se distinguen cuatro tipos de contextos:

- a. Contextos personales. Se relacionan con las actividades cotidianas que tienen relevancia personal directa e inmediata para el estudiante.
- b. Contextos educativos o laborales. Son situaciones a las cuales el estudiante pudiera enfrentarse en el ambiente escolar o en un entorno de trabajo. Incluye tareas que propone el profesor con fines exclusivamente didácticos.
- c. Contextos públicos. Son situaciones que surgen en la interacción diaria del individuo con el mundo externo.
- d. Contextos científicos. Son situaciones más abstractas con las cuales el estudiante está poco relacionado. Puede incluir un proceso tecnológico, una teoría o un problema eminentemente matemático.

El ciudadano que usa de manera efectiva el conocimiento matemático adquirido en una variedad de contextos necesita dominar ciertas competencias matemáticas específicas. Cuando se enfrenta a una tarea, el estudiante moviliza distintas capacidades y a diferentes niveles de ejecución. Cada una de las competencias que se trabajan en la evaluación PISA plantea distintos niveles de demanda cognitiva o diferentes niveles de profundidad. En el marco PISA (OECD, 2004), esta variable se denomina conjuntos de competencia (*competency clusters*). En este trabajo adoptamos la caracterización realizada por Rico y Lupiáñez (2008) que concretan esta variable en términos de niveles de complejidad de las tareas. En general, estas competencias están asociadas con tareas cuya dificultad va en ascenso; no obstante, existe superposición en la clasificación de las tareas dentro de cada categoría. Así, se distinguen los siguientes conjuntos:

- a. Reproducción. Las preguntas en este nivel requieren que el estudiante demuestre que domina el conocimiento aprendido. Son problemas que les

resultan familiares y se resuelven aplicando algoritmos o destrezas técnicas. Incluye los procesos de acceder (recordar, reproducir) e identificar.

- b. Conexión. Las preguntas en este nivel requieren que el estudiante demuestre que puede establecer relaciones entre distintos dominios matemáticos y que puede integrar información para resolver problemas que no son rutinarios pero que exigen que el estudiante se decida por una de entre varias estrategias de resolución. Incluye los procesos de aplicar, analizar y valorar.
- c. Reflexión. Las preguntas en este nivel son situaciones poco estructuradas que requieren que el estudiante comprenda, reflexione y use su creatividad para reconocer las matemáticas involucradas en el problema. Se exige que el estudiante analice, interprete y desarrolle sus propios modelos y estrategias y presente argumentos matemáticos, demostraciones y generalizaciones. Incluye los procesos de sintetizar, crear y juzgar.

5. Búsqueda y recopilación de las pruebas de diagnóstico

El acceso a las pruebas censales de evaluación se logró mediante una búsqueda sistemática en los portales de las Consejerías de Educación de cada comunidad. Limitamos la muestra a los ítems procedentes de las pruebas de cinco comunidades por lo cual, es una muestra por disponibilidad. Pese a varios intentos, dos comunidades no respondieron a la solicitud; tres comunidades sólo permitían acceso a la prueba aplicada en 4º de primaria bien porque no aplicaron prueba de competencia matemática en 2º de la ESO o porque no estaba disponible para divulgación; tres comunidades indicaron que la prueba no estaba disponible para divulgación porque sería usada para futuras aplicaciones; dos comunidades indicaron no haber evaluado competencia matemática en 2º de la ESO; una comunidad publicó sólo una parte de la prueba como ítems liberados; las pruebas de una comunidad serían liberadas en una fecha posterior y una comunidad sólo publicó ejemplos de ítems. En resumen, nuestro banco de información consta de cinco pruebas. Para garantizar la confidencialidad de las comunidades autónomas se asignó aleatoriamente un código a cada una de ellas: en lo sucesivo nos referiremos a ellas como comunidades A, B, C, D y E. En resumen, las pruebas disponibles contienen un total de 173 ítems distribuidos como se muestra en la siguiente tabla:

Tabla 1
Distribución del total de ítems por Comunidad Autónoma

Comunidad Autónoma	Número de ítems	Porcentaje
A	18	10,4
B	84	48,6
C	32	18,5
D	25	14,4
E	14	8,1

Puede observarse que no existe un patrón en lo que a número de ítems se refiere. La comunidad B incluyó la mayor cantidad de ítems con 84, representando esta cantidad cerca de la mitad del total de ítems en la muestra (48,6 %). Por su parte, la comunidad E contenía sólo 14 ítems, el menor número de ellos (8,1 %).

6. Análisis de ítems por Comunidad Autónoma y variable de tarea

De acuerdo con el modelo matemático PISA y el enfoque funcional del conocimiento matemático, se analizaron los 173 ítems disponibles en la muestra de cinco pruebas de diagnóstico. Para cada comunidad se evaluaron los ítems de acuerdo a las variables de tarea consideradas en el estudio: contexto, nivel de complejidad y contenido. Primero se analizaron los ítems de acuerdo al contexto. La siguiente tabla muestra la distribución porcentual de los ítems por comunidad y contexto.

Tabla 2
Porcentaje de los ítems por Comunidad Autónoma y contexto

Contexto	Comunidad Autónoma					Total
	A	B	C	D	E	
Personal	27,8	1,2	9,4	52,0	64,3	17,9
Educativo/Laboral	33,3	19,0	21,9	8,0	0,0	17,9
Público	38,9	69,0	59,4	40,0	35,7	57,2
Científico	0,0	10,7	9,4	0,0	0,0	6,9

En la distribución del total de ítems por el contexto se observa, globalmente, un desequilibrio hacia los ítems de contexto público, cuyos enunciados representan más del 50% del total. Este resultado se mantiene en todas las comunidades excepto la comunidad E, cuyos ítems se inclinan al contexto personal. El 6,9% del total de ítems se clasificó en la categoría de contexto científico destacando que las pruebas de las comunidades A, D y E no incluyeron ítems para este contexto. La comunidad E tiene la mayor proporción de ítems en el contexto personal (64,3%) sin incluir ítems en los contextos educativo/laboral y científico. La comunidad A incluyó el mayor número de ítems en el contexto educativo/laboral (33,3%) y no incluyó ítems en el contexto científico. La comunidad con la mayor proporción de ítems en el contexto público es la comunidad B (69,0%) así como en el contexto científico (10,7%).

Al analizar los ítems respecto del contenido se encontró que la distribución del total de ítems en las categorías de esta variable es equilibrada, con leves diferencias (Tabla 3). La mayor frecuencia se observa en el contenido de incertidumbre (27,7%). Las clasificaciones de cantidad y espacio y forma exhiben igual porcentaje (25,4%). Aunque las diferencias por comunidades son apreciables, globalmente hay equilibrio entre los contenidos utilizados para la totalidad de los ítems.

Tabla 3
Porcentaje de los ítems por Comunidad Autónoma y contenido

Contenido	Comunidad Autónoma					Total
	A	B	C	D	E	
Cantidad	16,7	21,4	28,1	44,0	21,4	25,4
Cambio y relaciones	33,3	29,8	6,3	4,0	21,4	21,4
Espacio y forma	38,9	25,0	25,0	16,0	28,6	25,4
Incertidumbre	11,1	23,8	40,6	36,0	28,6	27,7

Se aprecia en la distribución que la comunidad D incluyó la mayor proporción de ítems en el contenido de cantidad (44,0%). En el contenido de cambio y relaciones destaca la comunidad A con un 33,3%. Esta comunidad también incluyó la mayor proporción de ítems en el contenido de espacio y forma (38,9%). En cuanto al contenido de incertidumbre, fue la comunidad C quien incluyó el mayor número de ítems en este contenido (40,6%). Tanto la comunidad C como la D incluyeron las proporciones más bajas de ítems en el contenido de espacio y forma: 6,3% y 4,0%, respectivamente. En la comunidad B se observa una distribución equilibrada en el contenido de los ítems que, igualmente, se observa en la comunidad E más no así en el resto de las comunidades. La distribución porcentual de los ítems de acuerdo al nivel de complejidad se presenta en la siguiente tabla:

Tabla 4
Porcentaje de los ítems por Comunidad Autónoma y nivel de complejidad

Complejidad	Comunidad Autónoma					Total
	A	B	C	D	E	
Reproducción	27,8	45,2	56,3	60,0	35,7	46,8
Conexión	61,1	46,4	34,4	32,0	64,3	45,1
Reflexión	11,1	8,3	9,4	8,0	0,0	8,1

Las observaciones respecto del nivel de complejidad arrojan que la mayor frecuencia ocurre en el nivel de reproducción con 46,8%, aunque el nivel de conexión es muy cercano (45,1%). Invariablemente, el nivel de reflexión se aleja considerablemente de los niveles de reproducción y conexión: sólo el 8,1% del total de ítems se clasificó en este nivel. Este resultado se observa de la misma manera en la distribución por comunidades. También en este caso se detecta un sesgo global, debido al bajo porcentaje de ítems de reflexión. En la distribución de los ítems por nivel de complejidad en cada Comunidad Autónoma se observa que las

comunidades C y D exhiben la proporción más alta de ítems en el nivel de reproducción: 60,0% y 56,3%, respectivamente. La mayor proporción de ítems en el nivel de conexión se registró en las comunidades E (64,3%) y A (61,1%). La comunidad B, por su parte, incluyó la mayor proporción de ítems en el nivel de conexión (46,4%) manteniendo una proporción parecida en el nivel de reproducción (45,2%). La comunidad A incluyó poco más de tres quintas partes (61,1%) de sus ítems en el nivel de conexión. Esta comunidad incluyó mayor proporción de ítems en el nivel de reflexión que las demás comunidades (11,1%). Aunque la proporción más baja de ítems se incluyó en este nivel, sólo la comunidad E no incluyó ítems de reflexión.

7. Fortalezas y deficiencias de los ítems

Los 173 ítems evaluados en este estudio presentan distintas fortalezas y deficiencias. Según Bell, Buckhardt y Swan (1992), se considera que una tarea que moviliza destrezas y habilidades de orden superior, como las enmarcadas en el modelo de PISA y el enfoque funcional del aprendizaje matemático, exhibe fortalezas si cumple una o más de las siguientes condiciones: las situaciones que presenta son relevantes al estudiante; la relevancia del contexto se mantiene durante el proceso de resolución; la tarea no es rutinaria; promueve que el estudiante se desplace del contexto del problema al mundo matemático necesario para resolverlo; moviliza potencialmente distintas competencias; estimula la toma de decisiones; es secuencial. Si no cumple las condiciones mencionadas y si, además, la tarea propuesta atiende estrictamente a los contenidos curriculares y no a las competencias, se considera que el ítem presenta debilidades. La tabla 5 resume las fortalezas identificadas en los ítems evaluados, la Tabla 6 presenta las deficiencias y la Tabla 7, las competencias que potencialmente movilizan las tareas.

Tabla 5
Porcentaje de ítems con fortalezas

Fortalezas	Comunidad Autónoma					Total	Porcentaje (%)
	A	B	C	D	E		
Contexto relevante	9	61	13	19	9	111	64,2
Tarea no rutinaria	11	39	14	14	9	87	50,3
Moviliza distintas capacidades	10	47	7	9	7	80	46,2
Tiene relevancia práctica	4	42	4	18	7	75	43,3
Permite la toma de decisiones	10	23	7	5	5	50	28,9
Tarea secuencial	9	19	2	9	8	47	27,2
Hay desplazamiento del contexto al mundo matemático	4	8	0	0	3	15	8,7

Se destaca en la tabla que cerca de dos terceras partes de los ítems evaluados presentan un contexto relevante y necesario para resolver el problema, mostrando esta fortaleza con la mayor frecuencia (64,2%). La mitad de los ítems presentan una

tarea que consideramos no rutinaria. Con proporciones cercanas al 50% encontramos tareas enmarcadas en la vida real con relevancia práctica para el estudiante y tareas que movilizan distintas capacidades. Por el contrario, sólo el 8,7% de los ítems permite que el estudiante se desplace del contexto del problema al mundo matemático necesario para resolverlo. Es decir, que apenas una décima parte de los ítems incluidos en las pruebas estimula al estudiante a iniciar el proceso de matematización.

Tabla 6
Porcentaje de ítems con debilidades

Debilidades	Comunidad Autónoma					Total	Porcentaje (%)
	A	B	C	D	E		
Tarea corta que suprime la toma de decisiones	5	35	19	19	10	88	50,9
Tarea rutinaria	6	30	19	11	11	77	44,5
No moviliza distintas capacidades	5	29	21	14	6	75	43,4
Carece de relevancia práctica	6	27	26	7	7	73	42,2
Atiende contenidos curriculares	8	22	12	14	3	59	34,1
Contexto irrelevante o camuflado	5	15	19	6	2	47	27,2

Respecto de las deficiencias identificadas en las pruebas, se observa que en la mitad de los ítems la tarea es corta y suprime la toma de decisiones. Muy cerca de este porcentaje encontramos tareas rutinarias que, aunque enmarcadas en la vida real, carecen de relevancia práctica para el estudiante y tareas que no movilizan distintas competencias. Se observa, además, que poco más de una cuarta parte de los ítems presentan tareas cuyo contexto está camuflado o es irrelevante e innecesario para resolver el problema. De otro lado, aproximadamente una tercera parte de los ítems presenta tareas que en lugar de movilizar competencias, atienden estrictamente contenidos curriculares.

Además de clasificar cada ítem de acuerdo a las variables de tarea, se estudió qué competencias debía eventualmente poner de manifiesto el estudiante para resolver el problema y se realizaron algunas observaciones generales del ítem. Por ser una variable de desempeño y no de tarea, la identificación de las competencias se realizó como un recurso de apoyo para caracterizar el ítem. Enmarcada en el modelo matemático de PISA, esta caracterización tiene en cuenta los ocho tipos descritos anteriormente.

En cuanto a las competencias potenciales que movilizan las tareas, se observa en la Tabla 7 que representar es la capacidad potencialmente movilizada con mayor frecuencia (41,6%). Se observa, además, que la competencia de comunicar es potencialmente movilizada en el 35,3% de los ítems. La competencia de modelizar

pudo haber sido movilizada sólo en el 4,0% de los ítems, representando la competencia con el menor porcentaje.

Tabla 7
Competencias potenciales que movilizan las tareas de los ítems

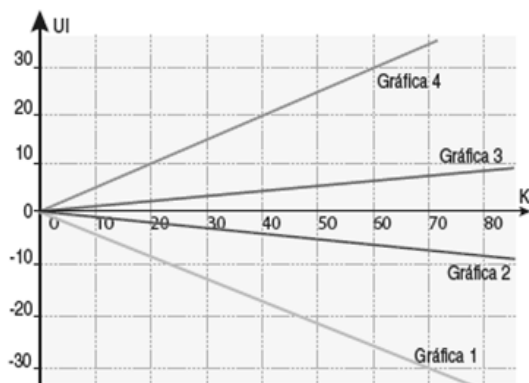
Competencias	Comunidad Autónoma					Total	Porcentaje (%)
	A	B	C	D	E		
Representar	7	38	7	13	7	72	41,6
Comunicar	6	35	5	8	7	61	35,3
Plantear y resolver problemas	7	11	3	0	3	24	13,9
Utilizar lenguaje simbólico, formal y técnico y las operaciones	7	20	2	0	2	31	17,9
Pensar y razonar	3	13	4	1	2	23	13,3
Argumentar	2	14	1	3	3	23	13,3
Modelizar	0	7	0	0	0	7	4,0

8. Ejemplo de ítems analizados

A continuación se presenta uno de los 173 ítems analizados y se identifican los valores de las variables que lo caracterizan.

La diabetes es una enfermedad que está aumentando de forma preocupante entre las personas jóvenes debido a los malos hábitos alimenticios y la obesidad. Consiste en que el cuerpo es incapaz de controlar los niveles de glucosa en sangre. El tratamiento más común es inyectar insulina diariamente en dosis adaptadas a cada persona. El cálculo de la dosis debe ser correcto, de lo contrario puede resultar perjudicial. La dosis diaria de la insulina "Lispro" es de 0,1 Unidades de Insulina (UI) por Kg de peso. ¿Qué gráfica representa la dosis de "Lispro" que hay que administrar por Kg de peso?

- A. Gráfica 1
- B. Gráfica 2
- C. Gráfica 3
- D. Gráfica 4



En nuestra investigación este ítem se clasificó en el contexto científico porque el estudiante está poco relacionado con la incidencia y prevalencia de enfermedades

como la diabetes; por su contenido corresponde a cambio y relaciones porque el estudiante interpreta datos presentados en la gráfica para establecer la relación entre las variables; por su nivel de complejidad es de conexión porque el estudiante debe reconocer la relación entre el peso y las unidades de insulina mediante la gráfica. Para resolver este problema se requiere que el estudiante: comprenda la información que se presenta, interprete la información para determinar la relación entre las variables, reconozca que existe una constante de proporcionalidad y analice la gráfica para determinar la solución. Esta tarea moviliza potencialmente las competencias: representar, pensar y razonar, comunicar. En este ítem se identificaron las siguientes fortalezas: el contexto es relevante, la tarea es no rutinaria, moviliza distintas competencias y permite la toma de decisiones basadas en información relevante (Caraballo, 2010).

9. Conclusiones

Con base en los objetivos que guiaron esta investigación, el marco de referencia que la sostiene y los resultados obtenidos, enunciaremos las siguientes conclusiones.

- Existen claras diferencias entre las Comunidades Autónomas respecto al acceso público a las pruebas aplicadas así como limitaciones en la transparencia del proceso.
- Los valores identificados en los ítems para las variables de tarea según definidas en el modelo matemático de PISA muestran que en la redacción de los ítems existen sesgos que favorecen el contexto público y al nivel de complejidad de reproducción.
- Los ítems redactados por las Comunidades Autónomas incluidas en la muestra mantienen un equilibrio en lo que a contenido se refiere.
- Puesto que en cada ítem se identificó un contexto, un contenido y un nivel de complejidad, variables de tarea definidas por el modelo matemático de PISA, los ítems aplicados se ajustan al modelo matemático de las evaluaciones de PISA.
- No obstante, por los sesgos y las deficiencias identificadas en ellas, las pruebas elaboradas por las Comunidades Autónomas para atender las evaluaciones de diagnóstico presentan limitaciones para satisfacer el propósito para el cual fueron diseñadas.

10. Recomendaciones

A la luz de las conclusiones derivadas de este estudio recomendamos a las Comunidades Autónomas que, para cumplir con el grado de ajuste adecuado a las evaluaciones PISA, revalúen el diseño de las pruebas a la luz de las variables definidas en su caracterización. Se sugiere que esta reevaluación sea orientada a subsanar los sesgos y las deficiencias identificados en esta investigación, mediante la inclusión equilibrada de ítems en las categorías de contexto, contenido y nivel de complejidad.

Bibliografía

Bell, A., Burkhardt, H. & Swan, M. (1992). Assessment of extended tasks. En R. Lesh & S. Lamon (Eds.). *Assessment of Authentic Performance in School*

- Mathematics*. Washington: American Association for the Advancement of Science.
- Boulmetis, J. & Dutwin, P. (2000). *The ABCs of Evaluation*. Jossey-Bass Publishers, San Francisco.
- Caraballo, R.M. (2010). *Análisis de los ítems de las pruebas de evaluación de diagnóstico en competencia matemática para el segundo curso de la Educación Secundaria Obligatoria en España, 2008-2009: un estudio exploratorio*. Trabajo de fin de máster sin publicar. Universidad de Granada.
- Instituto de Evaluación (2009). *Evaluación General de Diagnóstico 2009: Marco de la Evaluación*. Ministerio de Educación, Madrid.
- Lupiáñez, J. L. (2009). *Expectativas de aprendizaje y planificación curricular en un programa de formación inicial de profesores de Matemáticas de secundaria*. Tesis doctoral. Universidad de Granada.
- Lupiáñez, J. L. (2010). *Diseño y selección de tareas para el desarrollo de la competencia matemática*. Trabajo presentado en Matemáticas y Competencias Básicas, Oviedo.
- Ministerio de Educación y Ciencia (2006). Ley Orgánica de Educación. Boletín Oficial del Estado, BOE núm. 106. Autores, Madrid.
- OECD (2004). Learning for Tomorrow's World: First Results from PISA 2003.
- OECD (2009). Learning Mathematics for Life: A Perspective from PISA.
- Rico, L. (2006). "Marco teórico de evaluación en PISA sobre matemáticas y resolución de problemas". *Revista de Educación*, (Extraordinario 2006), 275-294.
- Rico, L. (2007). "La competencia matemática en PISA". *PNA*, 1(2), 47-66.
- Rico, L. & Lupiáñez, J. L. (2008). *Competencias matemáticas desde una perspectiva curricular*. Alianza Editorial, Madrid.

Rosa Marta Caraballo Caraballo: Profesora de Matemáticas a nivel universitario, editora y escritora de libros de texto de Matemáticas y consultora en el área educativa en Puerto Rico, cursa un doctorado en Didáctica de la Matemática en la Universidad de Granada, España. caraba@correo.ugr.es

Luis Rico Romero: Catedrático en la Universidad de Granada, Director del Departamento de Didáctica de la Matemática. Dirige el *Grupo de Investigación Didáctica de la Matemática*. *Pensamiento Numérico* del Plan Andaluz de Investigación, Desarrollo e Innovación, entre cuyas líneas se encuentran la formación del profesorado, la evaluación de programas de formación en matemáticas y el modelo de formación por competencias. lrico@ugr.es

José Luis Lupiáñez Gómez: Profesor del Departamento de Didáctica de la Matemática de la Universidad de Granada. Sus principales líneas de investigación son la formación de profesores de matemáticas de educación primaria y secundaria, la noción de competencia y el uso de nuevas tecnologías en el aula de matemáticas. lupi@ugr.es