

La competencia de ‘conocimiento e interacción con el mundo físico y natural’: Análisis de las pruebas de evaluación de diagnóstico de Andalucía

Monsalud Gallardo Gil, María José Mayorga Fernández, José Eduardo Sierra Nieto

Departamento de Didáctica y Organización Escolar. Facultad de Ciencias de la Educación. Universidad de Málaga. esierra@uma.es

[Recibido en julio de 2013, aceptado en febrero de 2014]

El presente artículo forma parte de una investigación centrada en la evaluación educativa de la competencia de Aprender cómo Aprender¹. Uno de los propósitos del estudio ha sido analizar el potencial de las evaluaciones externas y diagnósticas para estimar el grado de adquisición de las competencias básicas por parte de los estudiantes. En el marco de nuestro proyecto, se han establecido tres comisiones de investigadores² para abordar el análisis de las pruebas de Evaluación Diagnóstica de Andalucía (EDA), tal y como ya hicieron con las pruebas del programa PISA. El presente artículo recopila los hallazgos más sobresalientes de la investigación sobre las pruebas en el área específica de Ciencias, definida como ‘Competencia en el conocimiento y la interacción con el mundo físico y natural’. Tomando como matriz seis capacidades científicas (reproducción, aplicación, reflexión, transferencia, heurística y argumentación), se concluye que estas pruebas demandan principalmente capacidades científicas de baja complejidad con escasa presencia de capacidades de orden superior. Se especifica la metodología empleada, basada en análisis cualitativos y cuantitativos de las pruebas, distinguiendo las diferentes capacidades implicadas y el nivel de dificultad en el que aparecen. Se ofrecen los resultados más destacables tanto de la etapa de Primaria como de Secundaria en relación con la evaluación de competencias, que sometemos a discusión en el último apartado del artículo. También en este último apartado, se realiza una valoración de la distancia existente entre el concepto global de competencia propuesto en el informe DeSeCo (OCDE), el marco conceptual de la EDA, y lo que realmente parecen evaluar las pruebas a la luz de los resultados obtenidos.

Palabras clave: competencias básicas; competencia científica; educación; evaluación diagnóstica; evaluación externa.

‘Knowledge and interaction with the physical and natural world’ competence: Analysis of the diagnostic assessment tests Andalusia

This article is part of an investigation focusing on educational assessment competence Learn how to learn. One purpose of the study was to analyze the potential of external evaluations and diagnostic to estimate the degree of acquisition of basic skills by students. As part of our project, we have established three committees of researchers to address the analysis of the Diagnostic Evaluation tests of Andalusia, just as they did with the evidence from PISA. This article lists the most important findings of the research on testing in the specific area of science, defined as ‘Competence in knowledge and interaction with the physical and natural world’. Drawing on six matrix scientific capabilities (reproduction, application, reflection, transfer, heuristic and argument), we conclude that these tests mainly require low complexity scientific capabilities, with little presence of higher-order skills in them. It specifies the methodology, based on qualitative and quantitative analysis of the tests, distinguishing the different capacities involved and the difficulty level in which they appear. We offer the most

¹ La investigación que aquí presentamos se contextualiza en el Proyecto I+D, titulado: ‘La Evaluación Educativa de los Aprendizajes de Segundo Orden, Aprender cómo Aprender, Análisis de Proyectos Internacionales y Experimentación de Estrategias Alternativas’ (SEJ2007-66967), financiado por el Ministerio de Educación y dirigido por el profesor D. Ángel I. Pérez Gómez, a quienes agradecemos su inestimable ayuda para llevar a cabo dicho estudio. Nuestro más sincero agradecimiento va dirigido también al Servicio de Evaluación de la Junta de Andalucía, por facilitarnos el material necesario para llevar a cabo esta investigación.

² Grupo de Investigación del PAIDI. Plan Andaluz de Investigación, Desarrollo e Innovación): *Innovación e Investigación educativa en Andalucía*. URL: www.innova311.org

outstanding results of the stage both Primary and Secondary in relation to skills assessment, we submit for discussion in the last section of the article. As a final discussion section, we make an assessment of the gap between the overall competence concept proposed in the report DeSeCo (OECD), the conceptual framework itself Diagnostic Evaluation tests of Andalusia, and it really seems to assess the evidence in light of the results obtained.

Keywords: education; external assessment; diagnostic assessment; key competence; scientific competence.

Introducción

El propósito general del proyecto del que forma parte la investigación que aquí se presenta ha consistido en indagar en las posibilidades de identificar el grado de adquisición de *competencias de orden superior* o de *segundo orden* DeSeCo (OCDE, 2000) por parte de los estudiantes. A la luz de este marco general, y teniendo en cuenta la preocupación internacional y nacional que la evaluación de competencias está despertando (Consejos Europeos de Lisboa, 2000; Barcelona, 2002; Conferencia de Educación en España. MEC, 2007), se define como objetivo estudiar y valorar en qué medida las pruebas de PISA constituyen herramientas valiosas para evaluar competencias (Gallardo, *et al.*, 2010; Yus, *et al.*, 2013), en la línea de autores como Hernández (2006) y Gil y Vílchez (2006). De acuerdo con esto, se decidió ampliar el estudio para las pruebas de Evaluación de Diagnóstico en Andalucía (EDA), considerando que su propósito de fondo es “convertirse en el instrumento privilegiado para fomentar la mejora constante de los procesos de enseñanza-aprendizaje y la mejora del sistema educativo” (Junta de Andalucía, 2006). En este artículo se presentan los hallazgos derivados del análisis del ámbito científico de esta evaluación, concretamente de las pruebas centradas en la “Competencia en el conocimiento y la interacción con el mundo físico” (de aquí en adelante, CMFN), cuyos cuadernillos fueron facilitados por el Servicio de Evaluación de la Junta de Andalucía bajo un protocolo de confidencialidad. El propósito de fondo del estudio que presentamos consiste en valorar en qué medida las pruebas de EDA en el área de ciencias resultan valiosas para evaluar el grado de adquisición de competencias por parte de los estudiantes.

A lo largo del artículo se abordarán algunos de los dilemas que se presentaron en relación a las concepciones de competencia, se profundizará en la descripción del proceso seguido para el análisis de las pruebas y se presentarán las conclusiones que se derivan de éste.

Contexto internacional de evaluación externa: riesgos y posibilidades

En los últimos 15 años, las transformaciones en las economías mundiales, unidas a los cambios que las TIC han introducido en la vida productiva y cotidiana, vienen configurando un contexto denominado globalización o mundialización. Este contexto *macro* ha ido reclamando un debate alrededor de las necesidades educativas y formativas de los Estados, a la luz de las nuevas condiciones sociales y económicas. Necesidades que les llevan a tratar de mejorar los sistemas educativos según principios de calidad, accesibilidad y flexibilidad (Consejo Europeo de Estocolmo, 2001).

Una de las iniciativas que se deriva de esta agenda tiene que ver con el ‘interés’ por los resultados de los sistemas educativos, vehiculado a través de la evaluación externa del rendimiento de los estudiantes. Los argumentos sobre los que se sostiene dicha política de evaluación son: (1) el control democrático sobre la calidad de la educación como servicio público; (2) el interés por optimizar la inversión en educación, y garantizar una alta cualificación para la ciudadanía.

Según Rodríguez, (2011:206), el interés por estas pruebas tiene que ver con *“la posibilidad de controlar la educación bajo estándares comparativos entre países que tiene un currículum y tradiciones educativas distintas”*. Es este propósito, según la autora, el que viene generando el desarrollo de mecanismos externos que llevan la evaluación desde los *“contenidos cognoscitivos del currículum”* a la *“evaluación de competencias”*.

Por otra parte, como señalan Pérez y Soto (2011:172-173), existen una serie de riesgos o posibles desviaciones que se derivan del uso de evaluaciones externas estandarizadas:

- Tendencia a utilizar de manera normativa, no informativa, los indicadores obtenidos a través de test estandarizados.
- La excesiva especialización del lenguaje utilizada para la difusión de resultados, que impide o dificulta su debate crítico.
- La simplificación de los procesos educativos debido al trabajo con variables discretas.

Si bien los argumentos de los autores se refieren a los programas internacionales de evaluación como PISA, PIRLS o TIMSS, no queríamos pasar por alto estas consideraciones de cara a situar al lector en el panorama de las evaluaciones externas contemporáneas; contexto del que participa la EDA. Porque, como ha señalado Santos (2010), para que efectivamente una práctica de evaluación responda a principios de responsabilidad democrática hemos de cuidarnos de mantener vigente la pregunta por las finalidades.

La evaluación de diagnóstico en España y en Andalucía

Los compromisos adquiridos por España en materia de evaluaciones de diagnóstico se reflejan en el Consejo de Ministros de 2006 y, posteriormente, en la Conferencia de Educación en España de 2007. La propia LOE (2006) incorpora esta preocupación por las evaluaciones de diagnóstico. En los artículos 29 (para Educación Primaria) y 31 (para la ESO), se hace explícito el compromiso de realizar una evaluación de diagnóstico de las competencias básicas (CCBB) de los estudiantes, con un carácter formativo y orientador. A su vez, en el art. 144 se expone el marco de referencia para las administraciones educativas, de cara a realizar las evaluaciones de diagnóstico centradas en las CCBB del currículum.

En 2006 se aprobó en Andalucía la Orden por la que se *“regulaban las pruebas de la evaluación de diagnóstico y el procedimiento de aplicación en los centros docentes de Andalucía sostenidos con fondos públicos”*. Dicha orden, de 28 de junio, tomaba como referencia el marco general de la LOE, así como el marco internacional de recomendaciones acerca de la evaluación de los sistemas educativos, cuyo punto de inicio se situaba en el Consejo Europeo de Lisboa de 2000. En dicho Consejo se estableció el objetivo estratégico de tratar de *adaptar los sistemas educativos a las demandas de la sociedad del conocimiento y a la necesidad de mejorar el nivel y calidad del trabajo*. Las sucesivas concreciones de este objetivo tuvieron que ver, entre otros, con la definición de las nuevas destrezas básicas, y la elaboración del marco de las denominadas CCBB (Consejo Europeo de Barcelona 2002; Comisión Europea, 2004).

Es este marco de evaluación de competencias el que se recoge y adopta en el contexto andaluz. Así, en el preámbulo de dicha orden y en su art. 3, se especifica que el objeto de las evaluaciones de diagnóstico no será *“medir tasas brutas de adquisición de contenidos”*, sino *“comprobar el nivel de adquisición de las CCBB alcanzadas por el alumnado”*. Se insiste en el carácter formativo de los resultados, explicitándose que las pruebas no tendrán efecto alguno sobre la evaluación continua de los estudiantes.

Posteriormente, la Ley 17/2007 de 10 de diciembre de Educación de Andalucía (LEA), en su artículo 156, recoge las indicaciones de carácter general que realizó la LOE en materia de Evaluaciones de Diagnóstico, según las cuales se prevé realizar las pertinentes pruebas al finalizar el 2º ciclo de Educación Primaria y el 2º curso de ESO.

En la Orden de 28 de junio de 2006 y, más tarde, en el documento marco que elaboró la Consejería de Educación sobre el Modelo de EDA (2008), el rendimiento del alumnado es considerado un indicador de la calidad del sistema educativo. Benítez (2010:83) señala que las pruebas externas de rendimiento han de constituirse en un instrumento para la mejora de la educación a un triple nivel: (1) en cuanto al rendimiento de los estudiantes, (2) respecto de la práctica educativa y (3) en relación al propio sistema educativo; teniendo como propósito transversal la rendición de cuentas. Es por esto que las políticas de evaluación externa se han incorporado al Plan estratégico de Evaluación del Sistema Educativo andaluz 2009-2012.

Inicio de las pruebas y características principales del modelo de la EDA

La EDA comienza a aplicarse en el año 2006 para Primaria y Secundaria. Previamente se llevó a cabo una aplicación experimental de las pruebas para validarlas. En los cursos 2006/07 y 2007/08, las pruebas se centraron en las CCBB de “Comunicación lingüística” y “Matemáticas”. En el siguiente curso se incorporó la CMFN. Desde 2009, la Agencia Andaluza de Evaluación Educativa (AGAEVE) es la encargada de organizar la implementación de las pruebas y realizar el análisis de los resultados.

Según se recoge en el documento marco sobre la EDA (Junta de Andalucía, 2008:25), sus objetivos son:

- Mejorar los rendimientos educativos de todo el alumnado andaluz, independientemente de su punto de partida, a través de las medidas de mejora introducidas como consecuencia del proceso evaluador.
- Favorecer la cooperación e integración de los esfuerzos de la comunidad educativa para la mejora del rendimiento del alumnado.
- Conocer el grado de logro de los objetivos educativos así como la incidencia que pueden tener en ellos los procesos educativos y los factores contextuales.
- Potenciar modelos de evaluación formativa para la mejora del rendimiento.

Benítez (2010:84) enfatiza estos objetivos y los concreta en dos: ayudar a los centros a mejorar sus prácticas docentes y el rendimiento del alumnado; y facilitar a la Administración educativa información relevante para orientar la política educativa.

En cuanto a las características que definen el modelo de EDA, podemos hablar de las siguientes:

- Conjuga una evaluación censal con una aproximación muestral sobre un conjunto representativo de centros. Para la *evaluación censal* el propósito formativo está relacionado con “*proporcionar a los centros información relevante en la que basar las medidas necesarias para superar las diferencias existentes entre el nivel competencial que se espera que el alumnado desarrolle y que realmente ha alcanzado*” (Junta de Andalucía, 2008). Para el caso de la *evaluación muestral*, el propósito formativo es “*proporcionar criterios normativos para la valoración que haga cada centro y hará posible la elaboración de informes globales de cara a informar a la comunidad educativa acerca de los progresos conseguidos por el alumnado de los centros andaluces y de las medidas en ellos adoptadas para favorecerlos*” (Junta de Andalucía, Opus cit.). La

aproximación muestral permite, nuevamente según Benítez (2010:81), “*contrastar la objetividad y el rigor de los datos obtenidos en la evaluación censal*”.

- Se apoya en el sistema informático Séneca, para agilizar y favorecer la intervención de los distintos sectores implicados, y la rapidez en la obtención de la información.
- Las evaluaciones externas se focalizan en los niveles de competencia de los estudiantes, y no en las tasas brutas de adquisición de contenidos.
- Incorporan un cuestionario de contexto; el cual “*añade información acerca de los factores socioeconómicos y culturales del entorno familiar del alumnado, las características y medios de los centros escolares y las características personales del alumnado*” (Rodríguez, 2011:207).

La Evaluación de Diagnóstico está diseñada bajo un modelo de “pruebas objetivas”, con ciertas particularidades, sostenidas por los siguientes argumentos: la ausencia de subjetividad; la rapidez en su implementación y corrección; y la amplitud de muestra que permite cubrir. Sin embargo, en el propio marco teórico de las pruebas se reconoce la limitación de las mismas para medir aprendizajes complejos y profundos. Es por esto que se propone un modelo basado en “situaciones-problema”, el cual permite partir de casos que sirven como base para la interrogación, tratando de remitir a situaciones propias de la vida escolar y extraescolar del alumnado (Junta de Andalucía, 2008:37).

Las *situaciones-problema* están elaboradas sobre la base de uno o varios de los siguientes tipos de información:

- Textos escritos en los que se recogen informaciones diversa (p. e., anuncios o instrucciones).
- Textos orales breves, en soporte digital, que persiguen una finalidad comunicativa.
- Imágenes diversas (p. e., mapas, fotografías, gráficas).

Frente a cada uno de los casos, nos encontramos con diferentes tipos de preguntas:

- De respuesta cerrada (elección múltiple).
- Que exigen el desarrollo de procedimientos y la obtención de resultados.
- Abiertas que admiten respuestas diversas (pueden diferir de unos sujetos a otros).

CCBB y CMFN

La EDA adopta el modelo de CCBB que propone la LOE, traído a su vez del Informe DeSeCo (OCDE, 2002). En el citado documento marco sobre la EDA, se recoge la definición de competencia de DeSeCo, y se completa con la de Laisner (2000) (Tabla I); además de verse informada por el trabajo de Rychen y Tiana (2004).

La reflexión acerca de las CCBB forma parte de un debate mayor acerca del currículum y la reflexión acerca de los saberes que merece la pena poner al alcance de los estudiantes en el marco de la obligatoriedad, con el propósito de ayudarles a enfrentarse a los escenarios sociales contemporáneos (Junta de Andalucía, 2008:25-26).

Para el caso de la CMFN, ésta se sitúa en conceptos como “alfabetización científica”, “comprensión pública de la ciencia”, “educación CTS”. Según Acevedo, Vázquez y Manassero (2003:81), la alfabetización científica como finalidad educativa hunde sus raíces a mediados del siglo XX, si bien en la década de los 90 se produjo una revitalización y una reivindicación de la necesidad de abordarla como parte esencial de la educación básica y general de todas las personas.

Tabla 1. Definición de competencias.

DeSeCo (OCDE, 2000)	Capacidad de responder a demandas complejas y llevar a cabo tareas diversas de forma adecuada. Supone la combinación de habilidades prácticas, conocimientos, motivación, valores éticos, actitudes, emociones, y otros componentes sociales y de comportamiento que se movilizan conjuntamente para lograr una acción eficaz.
Laisner (2000)	Saber hacer complejo, resultado de la integración, movilización y adecuación de capacidades, conocimientos, actitudes y habilidades, utilizados eficazmente en situaciones que tengan un carácter común.

Fuente: elaboración propia a partir de Junta de Andalucía (2008)

PISA (OCDE, 2006) definió la Competencia Científica como *la capacidad de emplear el conocimiento científico para identificar problemas, adquirir nuevos conocimientos, explicar fenómenos científicos y extraer conclusiones basadas en pruebas sobre cuestiones relacionadas con la ciencia*. Los grados de dominio de esta competencia implican un conocimiento “de” la ciencia y “sobre” la ciencia. De la misma manera, en los marcos teóricos de PISA se definen 3 tipos de tareas o procesos científicos que están en la base de esta competencia: *identificar cuestiones científicas; explicar fenómenos; utilizar pruebas científicas*.

En el contexto español, la LOE (Capítulo III., art. 6.1), define el curriculum 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*.” El catálogo de CCBB que se establece comprende 8 competencias, entre las que está la “Competencia en el conocimiento y la interacción con el mundo físico.”

La LEA (art. 38) expone que “*El sistema educativo andaluz tiene como prioridad establecer las condiciones que permitan al alumnado alcanzar las CCBB establecidas para la enseñanza obligatoria*”.

El modelo de la EDA define la CMFN como la “*capacidad de utilizar el conocimiento científico, aplicar la metodología científica y ser consciente del papel que ejercen la ciencia y la tecnología en el desarrollo de la sociedad y el medio ambiente* (Junta de Andalucía, 2008:56). Esta competencia trata de dotar a la ciudadanía de los conocimientos, destrezas y actitudes básicos para interpretar desde un punto de vista científico la realidad en la que viven; así como participar con sentido crítico y responsable para lograr un mundo más justo desde el punto de vista social, más saludable y más sostenible (Junta de Andalucía, Opus cit.).

El citado Modelo Teórico recoge que esta competencia habrá de ser evaluada desde la conexión con el contexto próximo y habitual de los estudiantes, para conocer cómo los alumnos son capaces de poner en juego sus diferentes conocimientos y habilidades científicas ante situaciones problemáticas reales.

Como ocurre con PISA, para poder “traducir” la CMFN en pruebas estandarizadas, debe “sufrir” un proceso de desglose en capacidades y dimensiones de capacidades. Ese proceso se realiza poniendo en relación cada CCBB con los objetivos de cada etapa del sistema educativo. A partir de ahí se concretan las dimensiones que componen cada CCBB y, seguidamente, se desglosan los elementos de cada una (Junta de Andalucía, 2008).

Descripción del proceso metodológico

Llegado este punto cabe recordar el objetivo del estudio que presentamos. Como se señaló al inicio, nos propusimos analizar en qué medida las pruebas diagnósticas se tornan en un mecanismo valioso para la evaluación de competencias; en nuestro caso concreto, para la evaluación de la CMFN. Para ello se han analizado 42 ítems de Primaria y Secundaria respectivamente, correspondientes a los cuadernillos de los cursos 2008-2009 y 2009-2010, facilitados por el Servicio de Evaluación de la Junta de Andalucía para nuestra investigación bajo unos protocolos de confidencialidad de datos, que hemos tratado de respetar en la presentación del artículo.

El estudio se ha desarrollado en dos fases:

1. Análisis documental de los marcos teóricos de la EDA, y contraste con los de PISA y DeSeCo respecto de las concepciones de *competencia* y *competencia científica*.

2. Análisis de las pruebas:

La metodología utilizada se concreta en el análisis de contenido de las pruebas por grupos de expertos (cuatro en nuestro caso: dos Doctoras en Pedagogía, un investigador predoctoral en formación, y un Catedrático de Instituto, Doctor en Ciencias), que valoran independientemente los ítems de cada prueba en función de las categorías de análisis previamente establecidas. De esta forma, se establece un acuerdo entre los distintos miembros de la comisión, sometiéndose a discusión y al juicio de un nuevo experto las adscripciones de aquellos ítems que manifiestan importantes discrepancias en la valoración inicial.

En cuanto a las categorías de análisis, se han analizado tres niveles:

- Capacidades cognitivas predominantes requeridas para la resolución de los ítems de estas pruebas. Nuestras categorías de análisis son notablemente similares a las establecidas en los trabajos de Hopmann, Brinek y Retzl (2007) de la Universidad de Viena, quienes analizan y comparan las categorías cognitivas implícitas en las pruebas de PISA en relación a las pruebas diagnósticas realizadas en Francia, Alemania y Finlandia. Estos autores, siguiendo las revisiones críticas de la taxonomía de Bloom elaboradas por Anderson y Krathwohl (2001)³, y que han supuesto una reconstrucción sustancial de la misma para introducir la primacía de los procesos así como el fortalecimiento de la flexibilidad e interacción sobre la linealidad mecánica de la primera taxonomía, establecen las siguientes categorías fundamentales:

a) Reconocer y recordar

b) Comprender

³ La revisión crítica de Anderson y Krathwohl (2001) establece las siguientes dimensiones:

1. *Reproducción*, recuerdo y reconocimiento.

2. *Comprensión*: establecer el significado de mensajes instructivos sencillos y directos en diferentes lenguajes. Ello incluye interpretación, ejemplificación, clasificación, resumen, inferencia, comparación y explicación.

3. *Aplicación*: llevar a cabo procedimientos sencillos en una situación concreta. Ello incluye la ejecución y la implementación de rutinas, fórmulas y procedimientos sencillos y de aplicación mecánica.

4. *Análisis*: descomponer un todo en diferentes partes. Incluye diferenciación, organización, atribución.

5. *Evaluación*: elaboración de juicios basados en criterios. Incluye comprobación y crítica.

6. *Creación*: reunir elementos en nuevos conjuntos y en productos originales. Incluye síntesis, generación, planificación y producción.

Sobre esta revisión sustancial de la taxonomía de Bloom, Norman Webb (Hess, 2005; 2006) elabora una propuesta más funcional y dinámica de cinco categorías de capacidades y cuatro dimensiones de conocimiento:

- *Categoría de capacidades*: Recordar, Comprender, Aplicar, Analizar, Evaluar y Crear.
- *Dimensiones de conocimiento*: Factual, Conceptual, Procedimental, Metacognición.

- c) Aplicar
- d) Analizar
- e) Crear
- f) Evaluar

Reconocer, recordar, comprender el sentido de términos y proposiciones sencillas y aplicar fórmulas o procedimientos sencillos de forma mecánica, se consideran cualidades cognitivas de orden inferior o de primer orden; mientras que el análisis, la transferencia compleja, la evaluación y la creación se consideran cualidades cognitivas de orden superior o de segundo orden.

Apoyándonos en estos planteamientos, proponemos, para su uso flexible en esta investigación, la siguiente matriz de categorización de capacidades, que consideramos se adapta mejor a los planteamientos de las pruebas de PISA y de la EDA:

TABLA 2. Capacidades predominantes.

CAPACIDADES		DESCRIPCIÓN
REPRODUCCIÓN		Repetir de forma mecánica y memorística los conocimientos científicos
APLICACIÓN		Emplear los conocimientos científicos aprendidos previamente a situaciones sencillas y/o conocidas
COMPRENSIÓN - REFLEXIÓN		Entender un fenómeno científico y reflexionar sobre los conocimientos aprendidos en esta área
2º orden	TRANSFERENCIA	Aplicar el conocimiento aprendido a nuevas situaciones, ‘conexionando’ ideas, conceptos o hechos científicos
	HEURÍSTICA	Diseño de un ‘plan’ para la solución de una situación científica, o la descripción de los pasos que son necesarios seguir para llegar a dicha solución
	COMUNICACIÓN - ARGUMENTACIÓN	Razonar de forma argumentada para explicar el fenómeno científico que se trate, comunicando la conclusión científica a la que se llegue a través del lenguaje escrito

Estas capacidades se analizan en función de su presencia o no en las pruebas analizadas, estableciendo tres niveles de predominancia, que van de 0 a 3, siendo el 0 “no presente”, el 1 “ligeramente presente”, el 2 “presente” y el 3 “muy presente”.

- Relevancia del contenido. Para este nivel se analizó el grado en que las *situaciones-problema* propuestas en las pruebas están adecuadamente contextualizadas en el entorno cercano de los estudiantes; esto es, si aludían a asuntos reconocibles como reales en la vida cotidiana de los estudiantes. En el marco de las CCBB y, concretamente para la CMFN, es fundamental que las tareas y actividades conecten con los intereses e inquietudes de los estudiantes, tratando de favorecer un *aprendizaje relevante*; y conectando así con el concepto de *aprendizaje situado*, clave en la adquisición de competencias.
- Formato de las pruebas. Se prestó atención tanto al lenguaje y el estilo discursivo empleado en la presentación de las pruebas, como al tipo de actividad que demandan

por parte de los estudiantes. En cuanto al lenguaje, se consideró si permitía al estudiante entender con claridad la situación-problema planteada. En lo que respecta al tipo de prueba, se consideró si se trataba de una actividad con respuesta cerrada (única respuesta correcta) o abierta (múltiple, sujeta a reflexión).

Una vez completado el análisis de las pruebas se realizó un volcado informático de la información a través de un formulario diseñado para tal efecto, lo que nos permitió un tratamiento cuantitativo de los datos haciendo uso de un programa informático de análisis estadístico.

Los resultados obtenidos fueron sometidos a una posterior evaluación interjueces, de manera que, tras haber analizado por separado las diferentes pruebas, el total de ellas es revisado de forma conjunta con el propósito de validar los resultados obtenidos en cada una de las categorías analíticas propuestas (predominancia de capacidades, relevancia del contenido y formato de las pruebas). Este proceso de revisión introduce un proceso de discusión que repercute en la fiabilidad y validez final de los hallazgos obtenidos.

Análisis de las pruebas. Resultados

Análisis cuantitativo de capacidades predominantes

Para una mejor comprensión de las capacidades predominantes en cada prueba, se ha decidido realizar el análisis en base al nivel educativo de referencia.

Educación Primaria

Analizando los cuadernillos de Educación Primaria podemos apreciar las capacidades que se encuentran presentes en el alumnado.

En la figura 1 se aprecia que la capacidad de reproducción no está presente en un 79% de los ítems analizados (Nivel 0). Mientras que en un 17% está muy presente (Nivel 3). La presencia de los niveles intermedios son prácticamente inexistentes.

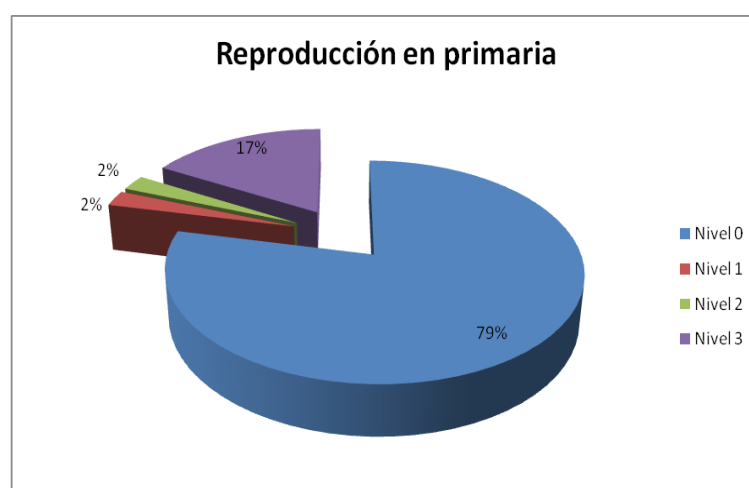


Figura 1. Capacidad de Reproducción en Educación Primaria (Fuente: elaboración propia).

La capacidad de aplicación en las pruebas diagnósticas aplicadas en educación primaria no está presente en un 45% de los ítems (Nivel 0), y en un 36% de los ítems se encuentra muy presente en un Nivel 3, mientras que en un 14% de ellos está representada (ver figura 2).

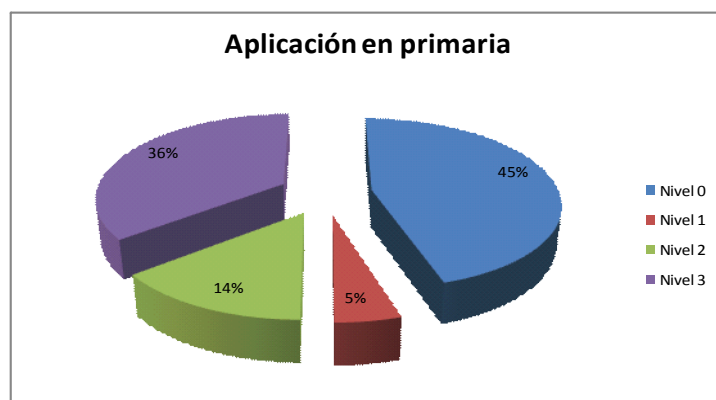


Figura 2. Capacidad de Aplicación (Fuente: Elaboración propia).

La capacidad comprensión/reflexión se encuentra muy presente (Nivel 3) en un 36% de los ítems, presente en un 24% (Nivel 2) y no está presente en un 31% (Nivel 0). Como se puede apreciar en la figura 3, en la mayoría de los ítems o no está presente o está algo presente. Respecto al nivel 1 se aprecia que su presencia es casi inexistente en un 9% de los ítems analizados.

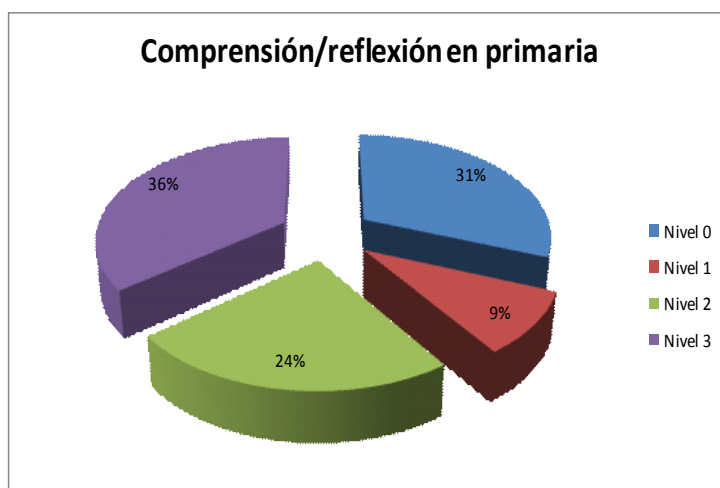


Figura 3. Capacidad de comprensión/reflexión en primaria (Fuente: Elaboración propia)

La capacidad de transferencia, como se observa en la figura 4 no está presente en la gran mayoría de los ítems (85% en un nivel 0). Solo en un 3% de los casos está muy presente (Nivel 3). Está ligeramente presente en un 5% de los casos y presente en un 7%.

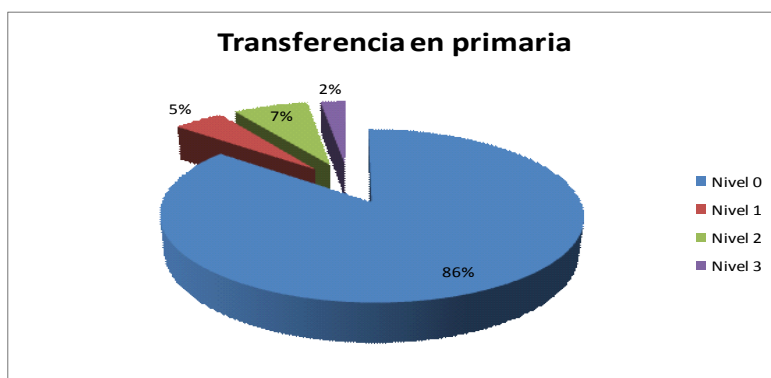


Figura 4. Capacidad de Transferencia en primaria (Fuente: Elaboración propia).

En la capacidad descrita en la figura 5 sucede igual que en la anterior, en prácticamente la totalidad de los ítems está ausente, es decir, no aparece en un 96% de los ítems analizados, mientras que sólo aparece en un Nivel 2 y en un Nivel 3 en el 2% de los ítems respectivamente.

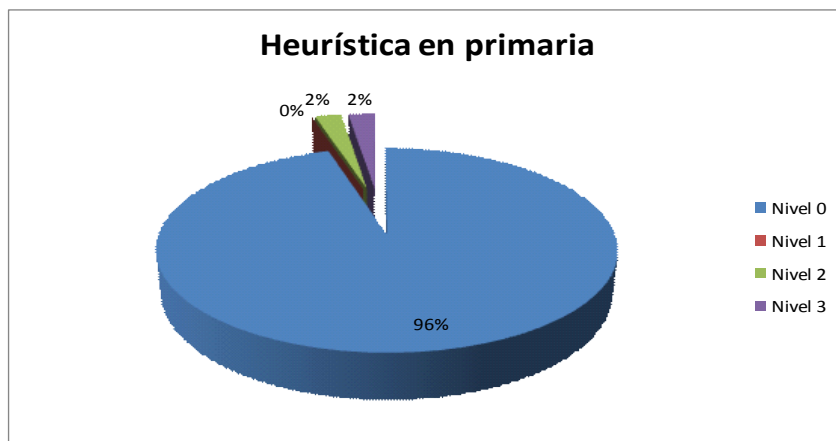


Figura 5. Capacidad de Heurística en primaria.

Finalmente, en un alto porcentaje de los ítems de los cuadernillos de Primaria, concretamente en el 81% de ellos, tampoco se aprecia esta capacidad. Y en el caso de hacerlo su presencia no es muy elevada, en un 5% de los casos. Sólo se aprecia que está muy presente en un 7% de los ítems. En un alto porcentaje de los ítems de los cuadernillos de Primaria, concretamente en el 81% de ellos, tampoco se aprecia esta capacidad. Y en el caso de hacerlo su presencia no es muy elevada, en un 5% de los casos. Sólo se aprecia que está muy presente en un 7% de los ítems.



Figura 6. Capacidad de Comunicación/argumentación en primaria (Fuente: Elaboración propia).

Educación Secundaria

A continuación se analizan las capacidades predominantes en los cuadernillos de Educación Secundaria Obligatoria (ESO).

En ESO la capacidad de reproducción está muy presente en el 43% de los ítems analizados, y bastante presente en el 15%, lo cual indica que un alto porcentaje de ítems implica reproducir contenidos (Figura 7).

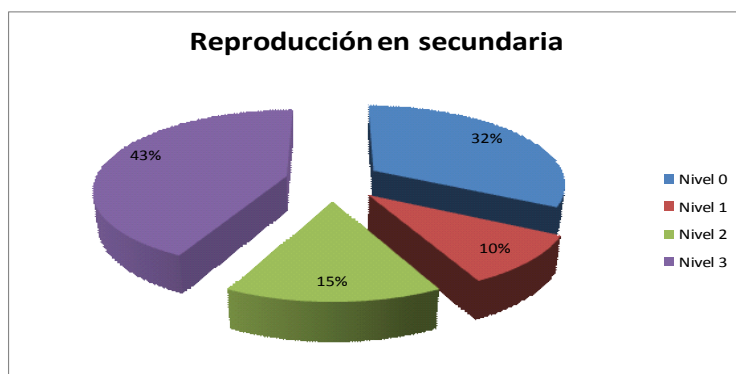


Figura 7. Capacidad de Reproducción en secundaria (Fuente: Elaboración propia).

En la capacidad mostrada en la figura 8 los valores obtenidos en los niveles 0 y 1 están igualados, lo cual indica que la capacidad de aplicación no está presente o su presencia no es muy evidente (60% de los ítems). Solo en un 8% de los ítems existe una presencia evidente de la capacidad de aplicación.

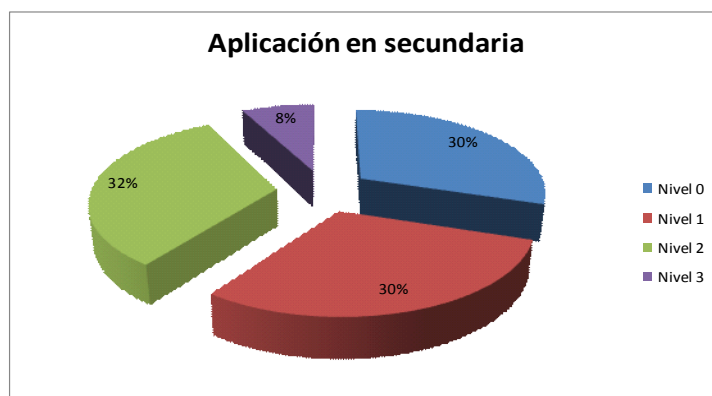


Figura 8. Capacidad de Aplicación en secundaria (Fuente: Elaboración propia).

La capacidad de comprensión reflexión está presente en un 12% de los ítems, mientras que en un 56% de ellos o no está presente o está presente en un nivel muy básico (Nivel 1).

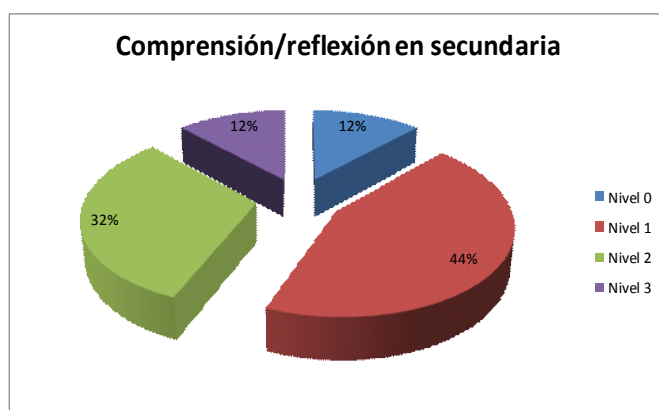


Figura 9. Capacidad de comprensión/reflexión en secundaria (Fuente: Elaboración propia).

La capacidad de transferencia no está presente en un nivel 3 en ninguno de los ítems, se encuentra ligeramente presente en un 5% de los ítems, presente en un 8% y en un nivel 0 un 87%, lo cual indica que esta capacidad no está presente en el cuadernillo analizado, y cuando aparece lo hace en un nivel básico (Figura 10).

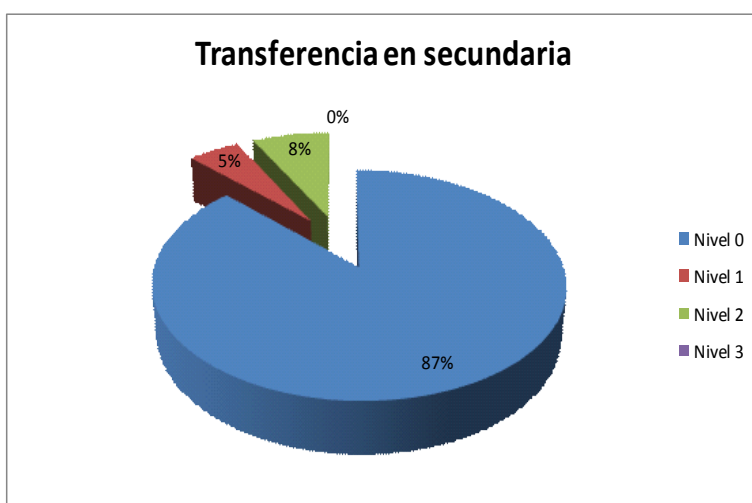


Figura 10. Capacidad de transferencia en secundaria (Fuente: Elaboración propia).

Como se puede apreciar en la figura 11, la capacidad de heurística es inexistente en los cuadernillos de ESO, en ningún ítem está presente.

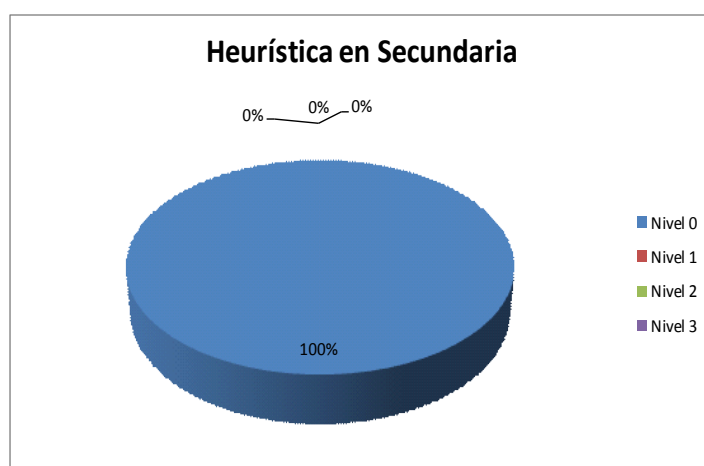


Figura 11. Capacidad de heurística en secundaria.

La capacidad de comunicación/argumentación está ausente en un 62% de los ítems, sólo está muy presente en un 8% de ellos (Nivel 3). Y en el 22% de los casos está relativamente presente, en un nivel 1.



Figura 12. Capacidad de comunicación/argumentación en secundaria (Fuente: Elaboración propia).

A grandes rasgos se puede afirmar que tanto en Primaria como en Secundaria la capacidad de reproducción está muy presente en un 30% de los ítems, la capacidad de aplicación en un 23%, la capacidad de comprensión/reflexión en un 10% de ellos, la capacidad de transferencia en un 1%, la capacidad de heurística en un 2% y, por último, la capacidad de comunicación/argumentación en un 8%. Con ello, se puede afirmar que las capacidades de orden superior están poco presentes en los cuadernillos analizados, tanto en Educación Primaria como en Educación Secundaria Obligatoria. Son las capacidades de orden inferior las que se encuentran de manera más evidente.

Relevancia del contenido

Educación Primaria

De los 24 ítems analizados del curso 2008-2009, encontramos que un 83'3% sí ofrece un contenido relevante; por su parte, de los 18 ítems de 2009-2010, encontramos relevancia del contenido en un 72'2%.

Tabla 3. Relevancia del contenido en Educación Primaria.

PRIMARIA		
	RELEVANCIA DEL CONTENIDO	
	SI	NO
2008-09	20	4
%	83.3	16.7
2009-10	13	5
%	72.2	27.8
% TOTAL ETAPA	78,6	21,4

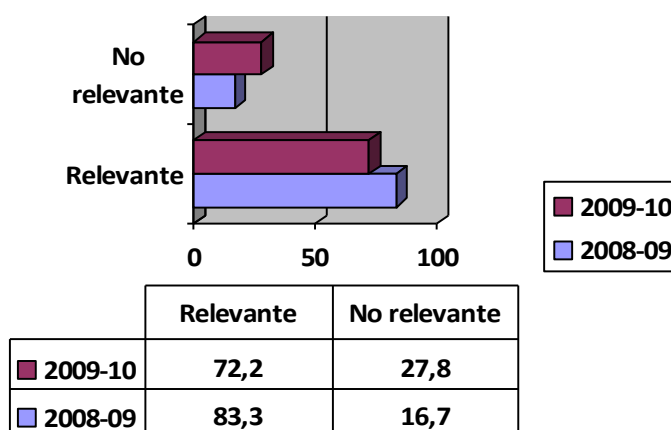


Figura 13. Relevancia del contenido en Educación Primaria (Fuente: Elaboración propia).

En suma, un 78'6% de los 42 ítems analizados sí presenta un contenido relevante frente a un 21'4% que no lo ofrece. Sin embargo, debemos matizar al respecto que nos hemos encontrado con la paradoja de valorar un contenido como potencialmente relevante aunque dicha relevancia quede finalmente muy comprometida debido al formato académico-tradicional de la pregunta, que resulta artificial y resta relevancia al contenido. Este hecho suele ocurrir en ítems en los que la actividad que se demanda al estudiante es sencilla, destacando los siguientes ejemplos: rellenar un esquema o un recuadro con los nombres, completar frases, "copiar y pegar", indicar con una x (señalar), unir con flechas o poner un número (relacionar),

seleccionar la respuesta correcta de las ya dadas, identificar verdadero/falso, ordenar, clasificar o agrupar. En el apartado de *formato* de las pruebas se profundizará en este análisis.

Educación Secundaria

En lo referente a los 22 ítems analizados para el curso 08-09, encontramos una diferencia de 8 puntos entre las consideradas como relevantes (45,8%) y las no relevantes (54,2%). Una diferencia que se incrementa hasta los 22 puntos en los 18 ítems analizados para el curso 09-10; donde nos encontramos unos porcentajes del 38,9% para los ítems considerados como relevantes, y de un 61,1% para los no relevantes.

Tabla 4. Relevancia del contenido en Secundaria.

SECUNDARIA		
	RELEVANCIA DEL CONTENIDO	
	SI	NO
2008-09	11	13
%	45.8	54.2
2009-10	7	11
%	38,9	61.1
% TOTAL ETAPA	42,85	57,14

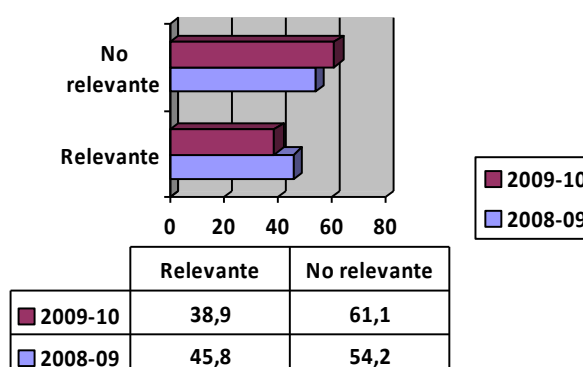


Figura 14. Relevancia del contenido en Secundaria (Fuente:Elaboración propia).

El análisis cualitativo pormenorizado de cada actividad nos desvela que la predominancia de actividades no-relevantes parece deberse fundamentalmente a dos razones. Por un lado, nos encontramos con que las situaciones-problema planteadas se refieren a problemáticas que escapan a la realidad cotidiana de los estudiantes. Por ejemplo, la actividad 23 (cuadernillo 08-09), cuyo propósito es que los estudiantes identifiquen 4 argumentos (de entre una lista de 10) a favor de la sostenibilidad del uso de placas solares; sin embargo, para contextualizar la pregunta recurren a un contexto internacional (un proyecto de electrificación de centros de salud con energía solar fotovoltaica en Cuba). Otro ejemplo de baja relación con el contexto próximo lo encontramos en la actividad 15 (cuadernillo 09-10), en la que se pregunta acerca de las consecuencias de la lluvia ácida en relación al p^H del agua y la existencia de peces, a partir de la interpretación de una gráfica. Un fenómeno como este es poco apropiado dado que escapa al contexto andaluz en el que se implementan las pruebas.

Por otra parte, nos encontramos con que las actividades se presentan de un modo académico, esto es, como un ejercicio típicamente escolar; algo también señalado para la etapa de Educación Primaria, y que va ligado al *formato* de las pruebas. Por ejemplo, nos encontramos

con actividades que se resuelven únicamente completando una tabla con datos que aparecen en el enunciado de la tarea (actividad 5, actividad 22, ambas del cuadernillo 08-09). O con la actividad 8 (cuadernillo 09-10), en la que se propone a los estudiantes que infieran las consecuencias de un experimento realizado con plantas a partir de un diagrama. Pese a que la actividad demandaría la experimentación científica, el contexto artificial de la misma le resta valor educativo.

Formato de las pruebas

Educación Primaria

Con respecto al formato, cabe indicarse que, en el cuadernillo correspondiente al curso 2008-2009, tan sólo un 37,5% de las preguntas presenta un formato abierto, frente a un 62,5% de cerradas. En el cuadernillo del curso 2009-2010, el porcentaje de respuestas abiertas es de 38,9%, siendo el total de respuestas abiertas 38,1 frente a un 61,9% de respuestas cerradas.

Tabla 5. Formato de las pruebas en Educación Primaria.

PRIMARIA				
	FORMATO			
	ABIERTO	CERRADO	LENGUAJE	
			CLARO	SENCILLO
2008-09	9	15	24	24
%	37,5	62,5	100	100
2009-10	7	11	18	18
%	38,9	61,1	100	100
% TOTAL ETAPA	38,1	61,9	100	100

Como se viene señalando, encontramos ítems de formato muy académico que recuerda a actividades clásicas de libro de texto. Así, por ejemplo, la pregunta 14 (cuadernillo 09-10) sobre infecciones (contenido potencialmente relevante) se limita a solicitar una X en la casilla correspondiente para indicar por dónde pueden entrar en nuestro cuerpo (nariz, boca, piel, genitales). Otro ejemplo muy esclarecedor de cómo el formato académico-tradicional resta potencialidad al contenido del ítem lo encontramos en la pregunta 6 (cuadernillo 08-09) en la que un contenido tan relevante como pueden ser las cuestiones de género queda desaprovechado al ser presentado de forma artificial; esto se debe a que el ítem pide marcar con una X allí donde corresponda (“Mujer, Hombre, No depende del sexo”) según las características presentadas (“sus órganos sexuales están fuera, tienen mamas, trabajan mejor, sus órganos sexuales están dentro, son más inteligentes, su voz es más grave, tienen más vello en el cuerpo”). O la pregunta 24 (cuadernillo 08-09) en la que el contenido sobre el consumo responsable queda artificial al limitarse la respuesta a la representación de un diagrama de barras con la información facilitada anteriormente sin más reflexión por parte del estudiante.

Con respecto al lenguaje utilizado, encontramos que es claro y sencillo prácticamente en el 100% de las preguntas. Además, todas ellas presentan un nivel de complejidad muy bajo, en ocasiones, mínimo, como es el caso de la pregunta 3 (cuadernillo 08-09), que puede llegar a resolverse “copiando y pegando” la información que aparece en el propio ítem.

Educación Secundaria

En lo que respecta a si los ítems proponen una solución abierta o cerrada, podemos señalar que, para el curso 08-09, encontramos un 62,5% con respuestas abiertas frente a un 37,5% cerradas. Para el curso 09-10, el porcentaje de respuestas cerradas aumentan hasta el 55,6%,

frente a un 44,4% de abiertas. Los resultados totales indican una predominancia de respuestas abiertas (54,8%), con una diferencia de casi 10 puntos sobre las cerradas (45,2%).

Tabla 6. Formato de pruebas en Secundaria.

SECUNDARIA				
	FORMATO			
	ABIERTO	CERRADO	LENGUAJE	
			CLARO	SENCILLO
2008-09	15	9	24	24
%	62,5	37,5	100	100
2009-10	8	10	18	18
%	44.4	55.6	100	100
% TOTAL ETAPA	54,8	45,2	100	100

Esto parece coincidir con el esfuerzo por diseñar ejercicios que privilegien la reflexión y la argumentación, aunque el simple hecho de que predominen tareas abiertas no parece garantizarlo, como se puede inferir del análisis de las capacidades predominantes.

En cuanto al lenguaje empleado, podemos concluir que para ambos cuadernillos analizados, el 100% de los ítems presenta las situaciones-problema con un lenguaje claro y sencillo. Sin embargo, creemos conveniente señalar que este hecho por sí mismo no aporta valor educativo a las pruebas, ya que es un nivel que cobra relevancia en la medida que acompaña, en una misma tarea, a otros niveles (por ejemplo: una actividad con un adecuado nivel de relevancia de contenido, que demanda capacidades de orden superior y que, además, es enunciada de manera clara y sencilla). La predominancia de pruebas poco relevantes, pese a contar con un enunciado y un planteamiento entendible, no poseen valor educativo. Además, por lo que hemos podido comprobar, esta sencillez significa “simpleza”, en tanto en cuanto un planteamiento potencialmente relevante es enfocado con una serie de preguntas con poca profundidad y que demanda en el estudiante capacidades de orden inferior.

Conclusiones y discusión

Las pruebas analizadas no parecen evaluar ‘competencias’ en el sentido complejo y global de DeSeCo (2002). En efecto, encontramos limitaciones de cara a la evaluación de la competencia científica en los términos que define el propio programa de la EDA –como un ‘saber hacer complejo’–.

¿Es posible traducir la compleja visión de la enseñanza y del aprendizaje que implican las competencias a un conjunto de pruebas de “papel y lápiz”? (Pérez y Soto, 2011). Rodríguez (2011:215) sostiene que es esa dificultad la que ha llevado a la Administración a elaborar unas pruebas que terminan por combinar “*el currículo por áreas con aspectos delimitados de las competencias*”. Crítica que amplía cuando señala cómo la excesiva fragmentación de competencias en dimensiones, y éstas en elementos de competencias, parece entrar en contradicción con la idea de competencia como sistemas complejos de comprensión y actuación (Pérez, 2007, 2008). Si bien, como es frecuente en el proceder científico, se comprendería una aproximación analítica para abordar un sistema complejo (en este caso la CMFN), el método empleado para su evaluación falla desde el momento en que luego se limita a verbalizar los resultados parciales, sin poder conectarlos entre sí en ese “todo” complejo que representa la CMFN; por lo que el resultado final es una lista de capacidades que por sí mismas no permiten evaluar el grado de adquisición de esta competencia.

Por otra parte, las pruebas se alejan de las capacidades de orden superior como son la transferencia, la heurística y la argumentación; haciendo, en definitiva, que la competencia pierda su sentido. El análisis de las capacidades da como resultado que la mayor parte se corresponden con las de orden más bajo (reproducción sobre todo en secundaria, aplicación), estando poco presente la comprensión profunda, casi totalmente ausente la capacidad de heurística y de transferencia, y también en muy bajo porcentaje la de argumentación. Si, como suele suceder con la aplicación de sistemas de evaluación externa, se evalúan principalmente capacidades de orden inferior, el sistema se adaptará a una enseñanza que fomente únicamente estas capacidades, con lo que no mejorará y no desarrollará competencia científica.

Como señalan Pérez y Soto (2011:175) al respecto de PISA, estas pruebas se limitan a la evaluación de aquellos componentes de las competencias que pueden expresarse a través de este tipo de instrumentos, como son los conocimientos como información y las capacidades. Ellos afirman que las pruebas y los test de evaluación externa no pueden constituir por sí mismas herramientas de evaluación educativa, ya que simplemente miden resultados y no diagnostican ni la naturaleza de los procesos ni identifican las causas o factores que intervienen en la construcción de ese determinado desarrollo.

Aunque el concepto de competencia está sujeto a interpretaciones diversas, es difícil sostener que el mero conocimiento científico sea una competencia, más allá de su valor académico tradicional; el “saber” no implica necesariamente un “saber hacer” y un “saber decir”, base de cualquier competencia. De hecho, en las pruebas analizadas no abundan precisamente preguntas de este tipo. Ante estos resultados, nos preguntamos si las Pruebas de Evaluación de Diagnóstico tienen por función evaluar la adquisición de competencias o si, como se deduce de muchas de las pruebas, trata de evaluar meramente aprendizajes académicos. Asimismo, nos preguntamos qué se entiende realmente por conocimiento científico o qué podría aportar el conocimiento científico para entender el papel de la ciencia y la tecnología en nuestras vidas y cómo son nuestros estilos de vida y sus consecuencias. Lo importante sería entonces comprender cómo el conocimiento científico ayuda a los chicos y chicas a entender sus estilos de vida y reorientarlos.

Por otro lado, gran parte de las preguntas planteadas no son relevantes para el alumnado de estas edades, por ser excesivamente académicas o porque no se vinculan ni con sus intereses ni con sus vivencias cotidianas. En ocasiones, la escasa relevancia del contenido tiene más que ver con la forma en que está formulada la prueba que con la relevancia en sí del tema, cayendo en la reproducción de información en lugar de la solución de problemas interesantes.

Además, en el formato de las pruebas encontramos algunos errores conceptuales o falta de precisión en su enunciado, lo que puede inducir a fomentar ideas erróneas en el alumnado, o bien, a confundirles en sus respuestas. En otros casos, se proporciona un enunciado muy largo, con información irrelevante.

De otro lado, algunos contenidos potencialmente relevantes están ‘desaprovechados’ por el formato de presentación. Se debería aprovechar todo el potencial crítico de las preguntas, en tanto que una tendencia llamativa de las pruebas parece ser la restricción de los problemas a temas puramente académicos, con escasa conexión con los temas transversales y cotidianos. Es cierto que algunas actividades plantean cuestiones de salud y medio ambiente, pero no profundizan en ellas, no plantean preguntas que revelen en profundidad la actitud de las y los estudiantes, ni plantean reflexiones o dilemas éticos ligados a la actividad científica.

En general, las preguntas tienen un índice de dificultad de moderado a bajo. Sólo excepcionalmente encontramos alguna actividad de alta dificultad.

Por ello, se puede afirmar que las pruebas de lápiz y papel difícilmente son una herramienta válida para evaluar competencias. Pese al esfuerzo por proponer enunciados claros y entendibles, y de demandar tareas con respuestas abiertas, predominan contenidos poco relevantes, así como la demanda de capacidades de orden inferior. Además de actividades que implican capacidades de baja complejidad y que desgranar en exceso la propia noción de competencia en tareas parciales y poco complejas.

Como alternativa, se debería potenciar una aproximación más holística e interdisciplinar a las competencias (Monereo y Pozo, 2007), porque no hay correspondencia entre el concepto de ésta que plantea DeSeCo y el propio programa de EDA (y la forma en que las pruebas analizadas propone evaluarlas). Se trataría de presentar pruebas que plantearan problemas familiares más próximos al alumnado, con un nivel de complejidad adecuado, que integrase todas las capacidades, no sólo científicas, sino también lingüísticas, matemáticas, etc., y otras transversales, y dar al alumnado oportunidades para que aplicase su competencia científica para resolverlo, y se debería permitir el acceso a todo tipo de material de consulta para responder a las preguntas, etc. Las pruebas no deben requerir sólo conocimiento verbal, sino capacidad para aplicar la lógica y los conceptos y procedimientos básicos de la ciencia que, como tales, no requieren memorización verbal sino comprensión significativa y funcional, siendo de gran importancia la oportunidad de razonar, argumentar y explicar.

Por tanto, si lo que se pretende es fomentar el desarrollo de competencia científica, es imprescindible que se fomenten capacidades científicas de 2º orden y que éstas sean priorizadas en los sistemas de Evaluación de Diagnóstico.

Asimismo, el desarrollo de la CMFN no debe vincularse al “recuerdo” de determinados hechos científicos, sino a saber aplicarlos de forma creativa. Esto convertiría el conocimiento científico factual en lo que debería ser: un recurso imprescindible para resolver problemas que requieran competencia científica y, por lo tanto, debería estar presente en el proceso de la evaluación de competencias.

Referencias bibliográficas

- Acevedo, J. A., Vázquez, A. y Manassero, M. A. (2003). El Papel de la educación CTS en una alfabetización científica y tecnológica para todas las personas. *Revista Electrónica de Enseñanza de las Ciencias*, 2(2), 80-111.
- Agencia Andaluza de Evaluación Educativa. *Plan estratégico de Evaluación General del Sistema Educativo Andaluz (2009-2012)*.
- Anderson, L. W. (Ed.), Krathwohl, D. R. (Ed.) (2001). *A taxonomy for learning teaching and assessing: A revision of Blooms taxonomy*. New York: Longman.
- Benítez, A. J. (2010). Evaluar competencias básicas: ¿una moda o una necesidad? *Cuadernos de pedagogía*, 397, 82-89.
- Comisión Europea (2004). *Competencias clave para un aprendizaje a lo largo de la vida. Un marco de referencia europeo*. Puesta en práctica del programa de trabajo “Educación y formación 2010.” Grupo de trabajo B: “Competencias clave.” [Comisión Europea, Dirección general de Educación y Cultura](#).
- Consejería de Educación (2008). *Modelo de Evaluación de Diagnóstico en Andalucía*. Sevilla.
- Dirección General de Ordenación y Evaluación Educativa (2008). Resolución de 30 de julio de 2008, de la Dirección General de Ordenación y Evaluación Educativa, por la que se desarrollan determinados aspectos de la Orden que se cita, por la que se regulan las

- Pruebas de la Evaluación de Diagnóstico y se establece el procedimiento de aplicación en los centros docentes de Andalucía sostenidos con fondos públicos. Boletín Oficial de la Junta de Andalucía, nº 161 (13 de agosto), 16-28.
- Dirección General de Ordenación y Evaluación Educativa (2011). *Resolución de 18 de noviembre de 2011, por la que se regulan las Pruebas de Evaluación de Diagnóstico 2011/2012*.
- Gallardo, G.; Fernández, M.; et al. (2010). PISA y la competencia científica. Un análisis de las pruebas de PISA en el área de Ciencias. *RELIEVE: Revista Electrónica de Investigación y Evaluación Educativa*, 16 (2), 1-17.
- Gil, D. y Vilches, A. (2006). ¿Cómo puede contribuir el proyecto PISA a la mejora de la enseñanza de las Ciencias (y de otras áreas de conocimiento)? *Revista de Educación*. Nº Extraordinario, 1, 295-331.
- Hargreaves, D. H. (2005). *Personalising learning 3: learning to learn & the new technologies*. London: Specialist Schools Trust.
- Hernández, F. (2006). El informe PISA: una oportunidad para replantear el sentido del aprender en la escuela secundaria. *Revista de Educación*, extraordinario PISA 1: 357-379.
- Hopmann, S. T., Brinek, G. y Retzl, M. (Eds.) (2007). *PISA zufolge PISA - PISA According to PISA. Hält PISA, was es verspricht? Does PISA Keep What It Promises? Reihe Schulpädagogik und Pädagogische Psychologie*. Bd.6. Germany: Wien Lit-Verlag.
- Junta de Andalucía (2006). Orden de 28 de junio de 2006, por la que se regulan las pruebas de la evaluación de diagnóstico y el procedimiento de aplicación en los centros docentes de Andalucía sostenidos con fondos públicos. Boletín Oficial de la Junta de Andalucía, nº 150 (04 de agosto), 13-16.
- Junta de Andalucía (2007). Ley 17/2007, de 10 de diciembre, de Educación de Andalucía. Boletín Oficial de la Junta de Andalucía, nº 252 (26 de diciembre), 5-36.
- Laisner, F. (2000). *Rèussir la formation par competentes*. Montreal: Guérin.
- Monereo, C. y Pozo, J. I. (2007). Competencias para (con) vivir con el siglo XXI. *Cuadernos de Pedagogía*, 370, 12-18.
- Organización para la Cooperación y el Desarrollo Económico (OCDE) (2002). *La definición y selección de competencias clave (DeSeCo)*. Resumen ejecutivo. Disponible en: www.OECD.org/edu/statistics/desecco
- Organización para la Cooperación y el Desarrollo Económico (OCDE) (2006). *PISA 2006. Marco de la evaluación. Conocimientos y habilidades en Ciencias, Matemáticas y Lectura*. Madrid: Santillana-MEC.
- Pérez, Á. I. (2007). *La naturaleza de las competencias básicas y sus aplicaciones pedagógicas*. Cuadernos de Educación de Cantabria. Gobierno de Cantabria. Consejería de Educación.
- Pérez, Á. I. (2008). ¿Competencias o pensamiento práctico? La construcción de los significados de representación y acción. En Gimeno, J. (Comp.), *Educación por competencias, ¿qué hay de nuevo?* (59-99) Madrid: Morata.
- Pérez, Á. I. y Soto, E. (2011). Luces y sombras de PISA. Evaluación de capacidades de razonamiento sobre problemas cotidianos. *Revista C&E Cultura y Educación*, 23 (2), 171-182.
- Rychen, D. S. y Tiana, A. (2004). *Developing Key Competencies in Education: Some Lessons From International And National Experience*. UNESCO Publishing / IBE.

- Rodríguez, C. (2011). Evaluación Para La Mejora. El Modelo De Evaluación e Diagnóstico de Andalucía. *Revista C&E Cultura y Educación*, 23 (2), 205-220.
- Santos, M. A. (2010). Los peligros de la evaluación. *Cuadernos de Pedagogía*, 397, 90-93.
- Yus, R.; Fernández, M.; Gallardo, M.; *et al.* (2013). La competencia científica y su evaluación. Análisis de las pruebas estandarizadas de PISA. En *Revista de Educación*, 360, enero-abril 2013, pp. 557-576.