

ELABORACIÓN DE ÍTEMS PARA PRUEBAS DE DIAGNÓSTICO DE MATEMÁTICAS (SECUNDARIA)

Claudia Lázaro del Pozo
lazaromc@unican.es
Universidad de Cantabria/IES Santa Clara, España

Núcleo temático: V-Recursos para la enseñanza y aprendizaje de las matemáticas

Modalidad: MC

Nivel educativo: Seleccionar uno de los siete niveles considerados

Palabras clave: evaluación, diagnóstico, PISA, TIMSS

Resumen

Siguiendo el modelo de las evaluaciones externas internacionales de diagnóstico podemos elaborar nuestras propias unidades de evaluación de Matemáticas en Secundaria.

El elemento principal de estudio va a ser el de las pruebas de Matemáticas de evaluaciones internacionales, como TIMSS y PISA, con el fin de explicar cómo se elaboran las unidades de evaluación y dar pautas que nos permitan crear nuestras propias pruebas de diagnóstico.

En primer lugar, se analizarán las características generales de las evaluaciones externas internacionales TIMSS y PISA, que miden resultados de alumnos en Matemáticas. Precisamente, sus marcos de evaluación proporcionan una guía de gran utilidad para conocer las características específicas de sus pruebas y nos indican cómo relacionar contenidos, competencias, criterios de evaluación, estándares de aprendizaje evaluables y referentes curriculares en el diseño de instrumentos de evaluación. Para finalizar, se planteará una propuesta de elaboración de ítems, a partir de la matriz de especificaciones que se habrá explicado previamente.

Evaluaciones externas de diagnóstico de Matemáticas

El tema de las evaluaciones externas, admite diferentes enfoques. El presente documento se centra en evaluaciones externas de Matemáticas con fines de diagnóstico, que siguen el modelo de estudios internacionales, como PISA. Este tipo de pruebas nos proporcionan información acerca del nivel competencial de los alumnos y del nivel de calidad del sistema educativo. Asimismo, aportan datos de interés que se deberían tener en cuenta en la toma de decisiones para la mejora de la calidad de un sistema educativo.

Por otro lado, el elemento principal de estudio va a ser el de las pruebas de Matemáticas de este tipo de evaluaciones, con el fin de explicar cómo se elaboran las unidades de evaluación y dar pautas que nos permitan crear nuestras propias pruebas de diagnóstico.

En primer lugar, recogeremos las características más generales de las evaluaciones externas internacionales, como PISA, que miden resultados de alumnos en Matemáticas. Precisamente, su marco de evaluación proporciona una guía de gran utilidad para conocer las características específicas de sus pruebas y nos indica cómo relacionar contenidos, competencias, criterios de evaluación, estándares de aprendizaje evaluables y referentes curriculares en el diseño de instrumentos de evaluación. Como anexo se plantea una propuesta de elaboración de ítems, a partir de la matriz de especificaciones que se habrá explicado previamente.

El programa PISA

PISAⁱ (*Programme for International Student Assessment*) es un estudio internacional de evaluación educativa de las competencias (lectora, matemática y científica) alcanzadas por los alumnos a la edad de 15 años. El estudio es impulsado por la Organización para la Cooperación y el Desarrollo Económico (OCDE). La primera aplicación del estudio PISA se realizó en el año 2000 y se repite cada tres años. La última aplicación de las pruebas PISA tuvo lugar en 2015.

PISA cuenta con una serie de procedimientos y estándares técnicos consensuados por los países participantes, que se aplican en el diseño de la prueba, en las traducciones, en el muestreo, en la recogida de datos, en los análisis y en la presentación de resultados, dirigidos a conseguir un alto nivel de fiabilidad y validez.

Para hacer posible el análisis a lo largo del tiempo y observar la evolución de los resultados, se incluyen ítems de anclaje (ítems no liberados: aquellos que no han sido difundidos públicamente) que se han empleado en ediciones anteriores, además de nuevas preguntas de la competencia principal en cada edición.

Una característica destacada de este estudio es la muestra elegida. PISA no trata de evaluar a los alumnos escolarizados en un determinado curso académico, sino a aquellos que estén comprendidos entre los 15 años y tres meses y los 16 años y dos meses en el momento en que se realiza la aplicación.

En 2015 40.000 alumnos en 1.000 centros españoles desarrollaron la prueba de PISA desde el 20 de abril hasta el 5 de junio. En esta edición participaron 75 países y 565.000 alumnos de todo el mundo y se han evaluado las competencias de Lectura, Matemáticas, Ciencias y, por primera vez, Resolución de problemas colaborativos.

La edición de 2012 se centró en matemáticas y se incluyó una prueba de resolución de problemas y otra sobre competencia financiera. La edición de 2015 se centró en la competencia científica y se realizó íntegramente por medios informáticos.

El marco de la evaluación está configurado de acuerdo a los siguientes elementos, que se recogen en la siguiente tabla:

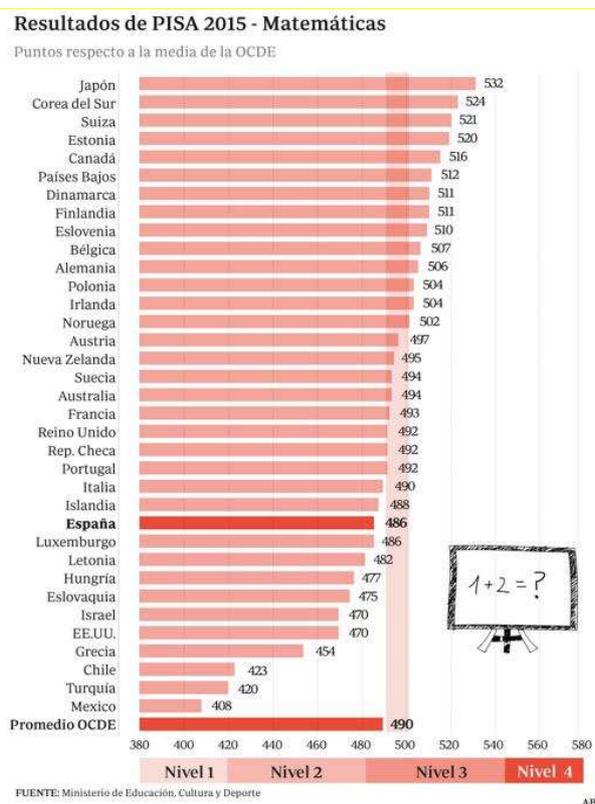
MARCO DE LA EVALUACIÓN PISA 2015 Matemáticas	
Contenidos	• Cantidad
	• Espacio y forma
	• Cambio y relaciones
	• Incertidumbre
Procesos cognitivos	• Formular (Transformar una situación real en algo susceptible de ser tratado de forma matemática).
	• Emplear (Aplicar razonamientos y herramientas matemáticas para obtener una solución matemática).
	• Interpretar (Reflexionar sobre los resultados matemáticos e interpretarlos en el contexto del problema).
Contextos y situaciones	• Personal
	• Educativo y laboral
	• Público
	• Científico

La primera aplicación del estudio PISA se realizó en el año 2000; las siguientes ediciones han sido en 2003, 2006, 2009, 2012 y 2015 proporcionando todas ellas gran cantidad de información que se recoge en el informe que con posterioridad se publica por cada una de las ediciones. El último informe se publicó en diciembre de 2016 y se refirió a los resultados de la aplicación realizada en 2015.

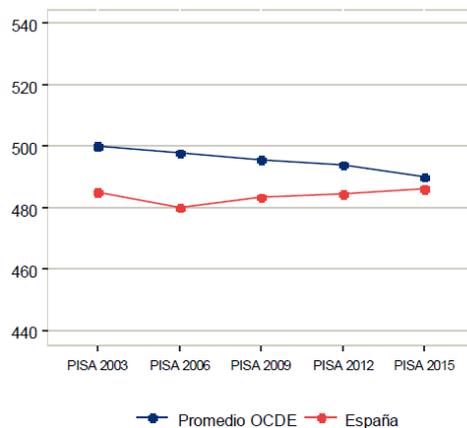
El estudio también recoge importante información de contexto de los propios alumnos y de los directores de los centros educativos a través de unos cuestionarios de contexto. Esta información contextual, puesta en relación con los resultados de las pruebas de rendimiento, ofrece una rica comparación de los rasgos que caracterizan a los sistemas educativos de los países participantes.

En las últimas ediciones de PISA no solo participó España como país, sino que también algunas comunidades autónomas optaron por ampliar su muestra, de modo que sus resultados tengan la precisión estadística suficiente como para poder ser comparadas entre sí y con los demás países participantes en PISA. En 2015, por primera vez las 17 comunidades autónomas españolas participaron en PISA con muestras representativas.

En la imagen siguiente se muestran los resultados en matemáticas en PISA 2015 en los países de la OCDE



En la siguiente gráfica se muestra la evolución de los resultados PISA en competencia matemática en el período 2003-2015



Fuente: MECD

Matriz de especificaciones de las pruebas de Matemáticas de PISA

La finalidad de una prueba depende de los objetivos a conseguir y, en función de estos, se diseñan los instrumentos que permiten medir, con la mayor precisión posible, aquello que queremos analizar. Para ello es necesario establecer los dominios de contenido, la vinculación con el currículo oficial, el alumnado objetivo, la duración de la prueba y las partes de las que consta, así como la población final de alumnado evaluada, a partir de la que se sacarán los resultados. Igualmente se hace imprescindible conocer cómo se desea analizar los resultados y las escalas de rendimiento que se emplearán.

Las evaluaciones internacionales se caracterizan por no estar vinculadas al currículo oficial de los países, fundamentalmente por las grandes diferencias que existen entre unos y otros, aunque con matices. Pretenden analizar la capacidad de respuesta y reacción frente a situaciones planteadas en las que el alumno, además de los conocimientos adquiridos en el aula, también debe poner en práctica otros que las diversas situaciones en las que se ha desarrollado su vida le han ido aportando (aprendizajes en la familia, el grupo de amigos, realización de actividades deportivas, vacaciones, etc.).

La matriz o tabla de especificaciones es la herramienta base que facilita y guía la construcción y la interpretación de pruebas. Los conceptos implicados son:

- Contenidos, declaraciones procedimentales que forman el cuerpo de enseñanza de una disciplina.
- Procesos o competencias, niveles de complejidad en la resolución de una tarea.

- Descriptores, formulaciones sintéticas de las competencias que habrán de ser medidas por los ítems que se elaboren.

El marco teórico de una evaluación internacional recoge toda la información necesaria para llevar a cabo desde el diseño de la prueba hasta su aplicación y de ello depende el análisis final que se pueda hacer del rendimiento.

A la hora de planificar una prueba hemos de tener claro la edad, o el nivel al que se le va a realizar, y cuál es el objetivo (diagnosticar, evaluar, etc). Hay que establecer el contenido de la prueba, las preguntas, las técnicas a incluir, el límite del tiempo, el número de ítems. Por otro lado hay que considerar los contenidos y competencias que se van a implementar y cómo se ponderarán los mismos. Estas y otras características son las que se definen en el marco teórico, pero sin duda la herramienta base que facilita y guía la construcción y la interpretación de pruebas es lo que denominamos matriz o tabla de especificaciones.

Para elaborar una matriz necesitamos establecer los contenidos, establecer los procesos cognitivos o grados de adquisición de la competencia, determinar los descriptores, determinar el peso relativo de cada casilla, donde se relacionan los contenidos con los procesos y elaborar ítems específicos para cada descriptor.

A continuación, se muestra un ejemplo para poder entender mejor lo que es una matriz de especificaciones, en la que se clasifican una serie de descriptores que apuntan biunívocamente a un grado de adquisición de la competencia y a la vez al bloque de conocimientos con el que se vincula.

Contenidos	Niveles cognitivos (grado de adquisición de la competencia)		
	Proceso I	Proceso II	Proceso III
Contenido A	Descriptores	Descriptores	Descriptores
Contenido B	Descriptores	Descriptores	Descriptores
Contenido C	Descriptores	Descriptores	Descriptores
Contenido D	Descriptores	Descriptores	Descriptores

Fuente de la imagen: Curso *Evaluaciones Externas Internacionales del Sistema Educativo*. Autor: INEE

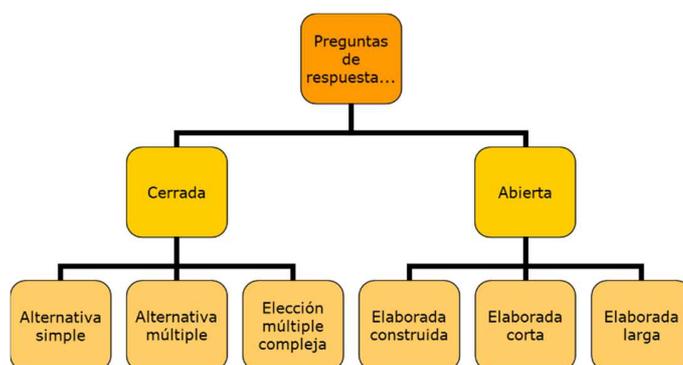
Estímulos y tipos de ítems

El estímulo o contexto de partida puede crearse a partir de varios elementos y, por tanto, podemos diseñarlos para que se adapten a las preguntas que vamos a realizar en referencia a los descriptores. En muchas ocasiones los estímulos son un texto de partida, como

generalmente ocurre en comprensión lectora, aunque se pueden incluir gráficos, esquemas, planos, medidas, cómics, etc.

Un ítem es una unidad de medida que consta de un reactivo (tronco) y una forma prescrita de respuesta. Su objetivo es obtener una conducta observable y claramente interpretable (debe ser capaz de discriminar). De la respuesta debe poder inferirse, con un grado suficiente de fiabilidad y validez, un nivel de dominio de la competencia. Para ello, un ítem ha de ser claro y conciso, no debe utilizarse para enseñar, debe proporcionar solo información precisa para responder, debe emplear formas positivas al redactar los textos siempre que sea posible y debe evitar las negaciones y, sobre todo, la doble negación.

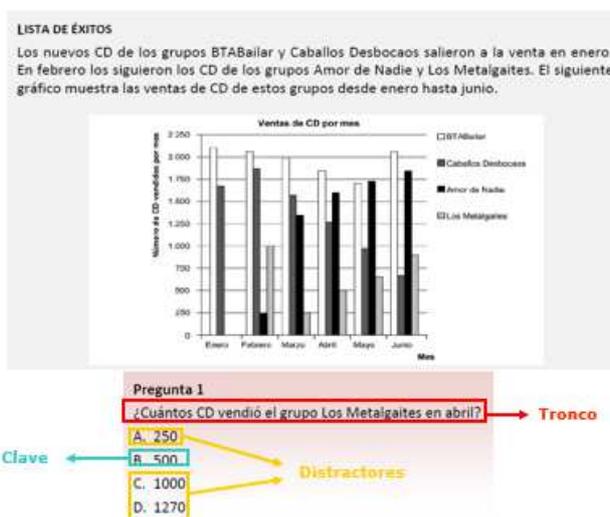
En la siguiente figura se recoge un esquema de la clasificación de los distintos tipos de ítems o preguntas que contiene una unidad de evaluación:



Los ítems de respuesta cerrada son los que muestran alternativas de respuestas, bien es decir, de respuesta alternativa simple (dicotómicas), combinan diferentes opciones de respuesta binaria (sí / no; verdadero / falso); o bien de respuesta de alternativa múltiple, que presentan varias alternativas (A, B, C, D). También, pueden ser de respuesta de elección múltiple compleja, combinando en un mismo ítem varias frases o alternativas con selección de respuesta binaria (sí / no; verdadero / falso).

Uno de los componentes que configuran un ítem de respuesta cerrada es el tronco o vástago, que es la parte inicial del ítem en la que se especifica la tarea: puede ser una pregunta, instrucciones, una frase incompleta, una situación de la vida real, etc. El otro componente de un ítem de respuesta cerrada son las opciones, es decir, todas las alternativas de respuesta de un ítem. De ellas, la clave es la respuesta correcta y los distractores son las opciones incorrectas.

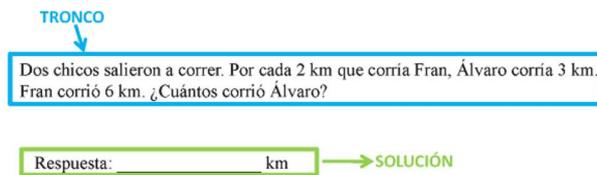
Se muestra a continuación un ejemplo de ítem de respuesta cerrada correspondiente a la prueba PISA



En los ítems de respuesta cerrada se deben tener en cuenta las siguientes consideraciones:

- La opción correcta debe ser claramente correcta.
- Los distractores que confunden a los expertos no son buenos.
- Se pueden emplear preconceptos erróneos como distractores.
- La dificultad del ítem se puede variar haciendo los distractores similares a la clave.
- No se debe abusar de opciones “todos los anteriores” – “ninguna de las anteriores”.
- No se debe repetir algo en todas las opciones
- Las opciones deben tener similar longitud, complejidad y estructura sintáctica.
- Con más opciones hay menor probabilidad de adivinación.
- Es recomendable solo una respuesta correcta.
- Evitar palabras en la clave que lleven a los que desconocen la solución a seleccionarla.
- Si el tronco es negativo, evitar opciones negativas.
- Se debe distribuir la posición de la clave al azar.

Los ítems de respuesta abierta se componen, también, del tronco o vástago y de la solución, que es la respuesta correcta. Pueden ser de respuesta elaborada construida, de respuesta elaborada corta (precisan de correctores con pormenorizados criterios de valoración), o de respuesta elaborada larga (precisan de correctores con pormenorizados criterios de valoración).



Fuente de la imagen: Curso *Evaluaciones Externas Internacionales del Sistema Educativo*. Autor: INEE.

Corrección de las pruebas

La corrección de una prueba también debe estar sometida a unos estándares preestablecidos que garanticen la homogeneidad en la corrección. Las preguntas de respuesta de elección múltiple son fáciles de codificar, pero las que permiten respuestas abiertas son más complejas. Algo que en ocasiones cuesta mucho esfuerzo a los correctores es no penalizar por cuestiones que en una prueba de aula se considerarían errores. Si de lo que el alumno expresa se sobrentiende que conoce lo que la pregunta indaga, se codificará como correcta o parcialmente correcta, en función de los códigos de cada ítem.

En realidad, el corrector o codificador no corrige la pregunta sino que compara la producción del alumno con unos criterios y en base a esta comparación asigna un determinado código. Para evitar los sesgos que se pueden producir por parte de un corrector a la hora de asignar los códigos, este proceso se somete a un control de calidad. Muchas de las respuestas dadas por los alumnos en los pilotajes son empleadas en el estudio principal como ejemplos de guía para facilitar el trabajo del codificador.

Referencias bibliográficas

- INEE (2016). Preguntas liberadas de PISA. <http://www.mecd.gob.es/inee/Preguntas-liberadas.html#PISA> Consultado 24/04/2017
- MECD (2016). *PISA 2015-Programa para la Evaluación Internacional de los Alumnos- INFORME ESPAÑOL*. Madrid: Secretaría General Técnica del MECD.
- OCDE (2016). *Marcos y pruebas de evaluación de PISA 2015*. Madrid: Secretaría General Técnica del MECD.

ANEXO: Elaboración de ítems para una prueba de diagnóstico de Matemáticas

Siguiendo el modelo de las evaluaciones externas internacionales de diagnóstico, podemos elaborar nuestras propias unidades de evaluación. Para ello, construiremos la correspondiente matriz de especificaciones en la que relacionaremos contenidos de aprendizaje y procesos cognitivos a través de un estímulo y una serie de ítems que nos van a permitir medir y describir el grado de cumplimiento de los diferentes criterios de evaluación y estándares de aprendizaje evaluables.

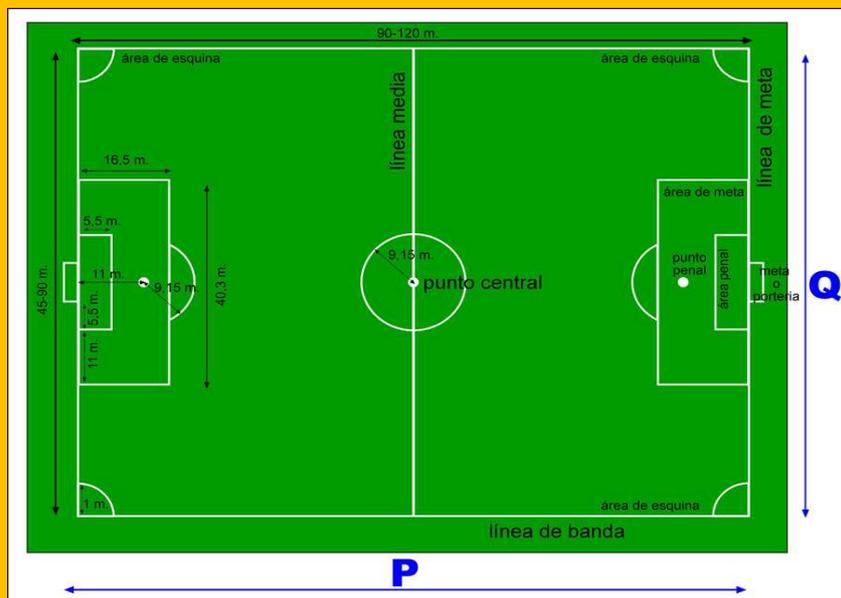
En el siguiente ejemplo se muestra una unidad de evaluación de este tipo diseñada para una evaluación de diagnóstico de nivel 2º de ESO (alumnos de 13-14 años)

Matriz de especificaciones						
Bloques contenidos	Reproducción		Conexión		Reflexión	
	Acceso e Identificación	Comprensión	Aplicación	Análisis y valoración	Síntesis y creación	Juicio y regulación
Números			Emplea adecuadamente los distintos tipos de números y sus operaciones para resolver problemas cotidianos contextualizados. (P5)	Identifica y discrimina relaciones de proporcionalidad numérica (como el factor de conversión o cálculo de porcentajes) y las emplea para resolver problemas en situaciones cotidianas. (P6)		
Álgebra	Calcula el valor de expresiones numéricas de distintos tipos de números mediante las operaciones elementales. (P1)	Identifica propiedades y leyes generales a partir del estudio de procesos numéricos recurrentes o cambiantes, las expresa mediante el lenguaje algebraico y las utiliza para hacer predicciones (P2)	Describe situaciones o enunciados que dependen de cantidades variables o desconocidas y secuencias lógicas o regularidades, mediante expresiones algebraicas, y opera con ellas. (P3 y P4)			

Geometría		Utiliza la escala para resolver problemas de la vida cotidiana sobre planos, mapas y otros contextos de semejanza (P7)			Analiza y comprende el enunciado de los problemas (datos, relaciones entre los datos, contexto del problema) (P8)	
-----------	--	--	--	--	---	--

CAMPO DE FÚTBOL

Las reglas del fútbol marcan que el terreno de juego debe ser rectangular, siendo su largo entre 90 y 120 m, y su ancho entre 45 y 90 m. Además, hay todo un conjunto de normas acerca de las medidas de los diferentes elementos de un campo de fútbol, que quedan recogidas en la siguiente imagen:



Como puedes observar, la recta que contiene la línea media del campo es un eje de simetría del terreno de juego.

1. Pedro y Juan encuentran en una revista las dimensiones del estadio de su equipo de fútbol, siendo $P = 100$ m y $Q = 70$ m. Como el área del terreno de juego viene dada por la expresión $P \cdot Q$, ¿cuánto mide el área del terreno de juego en este campo de fútbol?

- A. 7000 m
- B. 7000 m²
- C. 700 m
- D. 700 m²

Código 1: Respuesta B **Código 0:** Otras respuestas **Código 9:** Sin respuesta

2. ¿Cuál de las siguientes expresiones representa el perímetro del terreno de juego de un campo de fútbol?

- A. $2(P+Q)$

- B. $2P+Q$
- C. $P+Q$
- D. $P \cdot Q$

Código 1: Respuesta A **Código 0:** Otras respuestas **Código 9:** Sin respuesta

3. El área del terreno de juego viene dada por la expresión $P \cdot Q$. ¿Cuál es la superficie máxima del terreno de juego de un campo de fútbol reglamentario?

- A. 0,54 hm²
- B. 0,81 hm²
- C. 1,08 hm²
- D. 0,405 hm²

Código 1: Respuesta C **Código 0:** Otras respuestas **Código 9:** Sin respuesta

4. En el campo de fútbol del equipo de Pedro y Juan se pueden disputar partidos internacionales porque las medidas de su terreno de juego están comprendidas entre los siguientes límites:

	Mínima	Máxima
Longitud (P)	100 m	110 m
Anchura (Q)	64 m	75 m

¿Cuál es el valor mínimo que puede tener el perímetro del terreno de juego de un estadio en el que se disputen partidos internacionales?

Código 1: 328 m (o una medida equivalente en otras unidades). Es necesario especificar las unidades. **Código 0:** Otras respuestas **Código 9:** Sin respuesta

5. Pedro se ha colocado en el punto de penalti para entrenar su tiro a gol. ¿A qué distancia se encuentra el punto de penalti del borde del área de meta paralelo a la línea media y más lejano a la portería?

- A. 9,15 m
- B. 5,5 m
- C. 11 m
- D. Ninguna de las anteriores

Código 1: Respuesta B **Código 0:** Otras respuestas **Código 9:** Sin respuesta

6. Juan quiere mejorar su saque de córner. ¿Cuál es la razón entre el área del círculo central del terreno de juego y la de un área de esquina?

- A. $(9,15 / 4)^2$
- B. $9,152 / 4$
- C. $4 / 9,152$
- D. $9,152 \times 4$

Código 1: Respuesta D **Código 0:** Otras respuestas **Código 9:** Sin respuesta

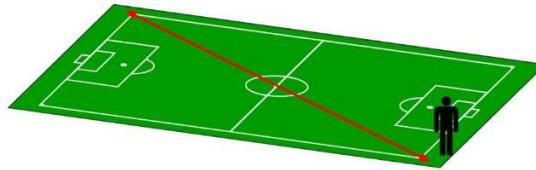
7. ¿Cómo se llama la figura que forma un área de esquina y cuál es el área total de las cuatro áreas de esquina?

- A. Sector circular, área= π m²

- B. Semicírculo, área= $\pi/4$ m²
- C. Semicírculo, área = π m²
- D. Sector circular, área= $\pi/4$ m²

Código 1: Respuesta D **Código 0:** Otras respuestas **Código 9:** Sin respuesta

8. Para entrenar Juan recorre corriendo la diagonal del terreno de juego, ida y vuelta, 30 veces todos los días. ¿Qué expresión representa la distancia que recorre el futbolista cada día?



Código 1: (una expresión algebraica correcta) **Código 0:** Otras respuestas
Código 9: Sin respuesta