

**EVALUACIÓN DE DIAGNÓSTICO EN
LA COMUNIDAD AUTONÓMA VASCA**

**CARACTERÍSTICAS DE
LAS PRUEBAS DE
RENDIMIENTO**

1. ¿Qué duración tendrán las sesiones de aplicación de las pruebas de evaluación diagnóstica?

Para decidir la duración de las sesiones en la que se aplicarán las pruebas se han tenido en cuenta dos aspectos claves:

- La edad del alumnado y su capacidad para mantener la concentración. Es necesario tratar de evitar el cansancio y la fatiga de las y los estudiantes, ya que el nivel de atención decae considerablemente transcurrido un cierto tiempo, por lo que las respuestas podrían ser irreflexivas y perderían su valor.
- La disponibilidad en cuanto a recursos necesarios y horario para la aplicación de las pruebas externas.

Un modelo razonable y ajustado a las características y cultura educativa de los niveles que se evaluarán no debe exceder los **dos días**, a lo largo de los cuales se incluirían la respuesta a las pruebas, la cumplimentación de un cuestionario y los descansos.

Para calcular el tiempo de dedicación real con el que se deberá contar para la aplicación de la prueba de evaluación de una competencia básica, es preciso tener en cuenta que se suelen necesitar unos cinco o diez minutos para presentar las características de cada prueba al alumnado (leer las instrucciones de respuesta, explicar la corrección de errores, visualizar algún ejemplo, crear un ambiente propicio y dar la oportunidad para aclarar dudas) y que, además, es preciso realizar descansos entre las pruebas (de cinco minutos entre prueba y prueba y uno más largo, de 20-30 minutos, en algún momento de la mañana).

Dada la diferente organización horaria que suelen tener los grupos de Educación Primaria y de ESO, a continuación se presentan dos propuestas específicas de jornadas de evaluación, correspondientes a cada una de los niveles en que se aplicarán las pruebas:

En **4º de Educación Primaria**, el tiempo máximo que se dedicará a cada sesión, es decir, a cada prueba no excederá los 45 minutos, en el caso de las pruebas de competencias no lingüísticas. Como el tiempo máximo disponible en algunos centros escolares es de tres horas en una mañana, las pruebas se distribuirían aproximadamente de la manera que se muestra en el cuadro de la página siguiente.

En esta distribución se ha tenido en cuenta la conveniencia de no evaluar en una misma mañana las competencias relacionadas con cada una de las lenguas, aunque éstas se evalúen mediante pruebas diferenciadas. El orden de evaluación de cada una de las competencias señaladas puede variar dependiendo de las necesidades de aplicación y de los horarios específicos de cada centro.

En **2º de la ESO**, el tiempo dedicado a la respuesta de cada prueba se amplía, dadas las características del alumnado, hasta 55-60 minutos por sesión en cada una de las competencias no lingüísticas. Además, también se amplía el módulo horario disponible, que se sitúa en torno a las tres horas y media cada uno de los dos días de aplicación, tal y como se muestra en el cuadro de la página siguiente.

4º de EDUCACIÓN PRIMARIA**PRIMER DÍA: ± 3 horas**

Organización y presentación de la evaluación al alumnado: 10'-15'

Competencia en comunicación lingüística: Euskara

Irakurmena: 25' + 5' = 30'
 Idazmena: 20' + 5' = 25'
 Entzumena: 15' + 5' = 20'
 Total: 1 hora 15'

Recreo: 30'

Competencia matemática 45' + 5' = 50'

Cuestionario: 15' + 5' = 20'

SEGUNDO DÍA: ± 2 horas 50'

Presentación de las pruebas del segundo día: 10'-15'

Competencia en comunicación lingüística: Lengua castellana

Comprensión lectora: 25' + 5' = 30'
 Expresión escrita: 20' + 5' = 25'
 Comprensión oral: 15' + 5' = 20'
 Total: 1 hora 15'

Recreo: 30'

Competencia científica, tecnológica y de la salud

45' + 5' = 50'

2º de E. SECUNDARIA OBLIGATORIA**PRIMER DÍA: ± 3 horas y ½**

Organización y presentación de la evaluación al alumnado: 10'-15'

Competencia en comunicación lingüística: Euskara

Irakurmena: 35' + 5' = 40'
 Idazmena: 25' + 5' = 30'
 Entzumena: 15' + 5' = 20'

Total: 1 hora 30'

Recreo: 20'-30'

Competencia matemática 55' + 5' = 60'

Cuestionario: 15' + 5' = 20'

SEGUNDO DÍA: ± 3 horas 15'

Presentación de las pruebas del segundo día: 10'-15'

Competencia en comunicación lingüística: Lengua castellana

Comprensión lectora: 35' + 5' = 40'
 Expresión escrita: 25' + 5' = 30'
 Comprensión oral: 10' + 5' = 20'

Total: 1 hora 30'

Recreo: 20'-30'

Competencia científica, tecnológica y de la salud

55' + 5' = 60'

2. ¿De cuántos cuadernos constará cada una de las pruebas?

Al tratarse de una evaluación externa de la que se quieren obtener resultados individuales, la prueba será básicamente común; es decir, el alumnado debe responder a los mismos ítems y en las mismas condiciones. Esto supone que sólo se contará con un único cuaderno que recogerá todas las cuestiones de la prueba. De esta forma se garantiza que todo el alumnado ha debido superar el mismo nivel de dificultad y se asegura la comparabilidad de los resultados.

La opción por un modelo de cuaderno único tiene como consecuencia que es menor el número de aspectos y dimensiones de cada competencia que se pueden medir.

En los cuadernos se mantendrán todos los años algunos ítems de anclaje. Se trata de ítems que hayan demostrado su fiabilidad, estabilidad y alta significatividad a la hora de evaluar un aspecto concreto de una competencia. La inclusión de este tipo de ítems permite realizar comparaciones longitudinales, es decir, interanuales y valorar la evolución dinámica de los resultados. La comparabilidad es interesante para los centros, que podrán comprobar el grado de eficacia de las medidas que han ido tomando y, por tanto, les ayudará en su proyecto de reflexión y autoaprendizaje; y, asimismo, es valioso para la administración educativa, que obtiene datos globales sobre el sistema y puede plantearse el análisis de buenas prácticas en los centros.

3. ¿Qué características formales tendrán los cuadernos?

Las pruebas se presentarán en soporte papel, con las excepciones que el propio contenido de algunas pruebas impone (por ejemplo, evaluar la comprensión oral exige registros en audio para ser reproducidos al alumnado en el momento de la aplicación).

Desde el punto de vista formal, los cuadernos serán fáciles de usar por el alumnado. Para las pruebas de lápiz y papel, se tendrá un especial cuidado en relación con el tamaño de la hoja, el tipo y el tamaño de la letra, la amplitud de los márgenes, los espacios para la respuesta o, entre otros aspectos, como el color empleado en algunos dibujos y gráficos.

Hasta este momento, en todas las evaluaciones llevadas a cabo el alumnado ha contestado a las cuestiones planteadas en el mismo cuadernillo y posteriormente se han introducido las respuestas en una base de datos y se han corregido los ítems abiertos en el mismo cuaderno. Sin embargo, no parece la opción más conveniente en una evaluación como esta con un tan alto número de alumnas y alumnos (alrededor de 35.000 cada curso), por lo se utilizarán **hojas de respuesta** separadas del cuadernillo que contenga la prueba por las siguientes razones:

- Reduce la cantidad de papel que es necesario utilizar ya que posibilita reutilizar los mismos cuadernillos de preguntas en tantas ocasiones como sea necesario.
- Es un procedimiento más cómodo en su uso, ya que ofrece la posibilidad de digitalización y facilita la corrección de las respuestas, especialmente cuando éstas corresponden a una prueba objetiva.

La opción por una hoja de respuestas separada del cuaderno exige una especial atención a su formato, sobretodo en el caso del alumnado de 4º de E. Primaria, que

probablemente esté poco o nada habituado a este tipo de instrumentos. En este sentido, por un lado, se cuidará formalmente la hoja de respuestas a fin de que sea clara, sencilla y atractiva para el alumnado que debe responder; y por otro, se propondrán procesos para que el alumnado conozca el mecanismo de respuesta y pueda probar algún modelo semejante antes de que tenga que hacer la prueba.

En los casos en que las respuestas impliquen la redacción de un texto, la elaboración de un dibujo o la elaboración de algoritmos de cálculo, es decir, tareas que generen registros de una cierta extensión, también serán recogidas directamente en la hoja de respuestas.

4. ¿Qué características tienen los ítems y a qué tipología responden?

La evaluación de competencias básicas en el alumnado de Educación Primaria y Secundaria Obligatoria requiere el empleo de instrumentos que incluyan ítems adecuados al tipo de competencias que han sido consideradas, que tengan en cuenta los contextos o situaciones definidos para que los sujetos demuestren su dominio y aplicación, y cuya administración resulte viable en el marco de una evaluación censal.

Las diferentes competencias objeto de evaluación exigen plantear diferentes tipos y formatos de ítems. No obstante, todos ellos tendrán en común su orientación hacia la evaluación de competencias básicas, sin centrarse exclusivamente en conocimientos y técnicas propios del currículo aplicado en cada contexto escolar. Las cuestiones, por tanto, basándose en los contenidos curriculares, tratan de evaluar las competencias de los sujetos para la aplicación de los mismos a situaciones escolares y especialmente a situaciones en las que habrá de desenvolverse en la vida real.

Para la elaboración de los ítems que integran las pruebas se han tomado en consideración tres referentes:

- El currículo y la descripción de las competencias realizada en el Decreto 175/2007 de enseñanzas básicas, junto con los documentos de desarrollo de cada una de ellas, de forma especial el apartado de indicadores de cada subcompetencia.
- En el caso de las pruebas que miden las competencias lingüísticas en ambas lenguas oficiales, euskara y castellano, las pruebas que han servido para medir los niveles B1 y B2 en euskara, tanto en cuanto al modelo y características de los ítems propuestos en las mismas, como en relación con el nivel planteado, que será coherente con los cursos en los que se aplica esta evaluación y con los niveles planteados en el citado Decreto 175/2007.
- La experiencia adquirida en la participación en las pruebas internacionales PISA (OCDE) y PIRLS (IEA), que buscan medir la capacidad del alumno/a para aplicar lo aprendido en el contexto escolar a la comprensión, análisis o resolución de problemas.

Algunas de las premisas a partir de las cuales se han formulado y construido los ítems son el entorno, los ámbitos de interés del alumnado evaluado, el contexto de la vida cotidiana, la funcionalidad, la resolución de problemas, la aplicación de lo aprendido, la autonomía para resolver problemas en situaciones distintas, los cambios continuos en la sociedad actual, la diversidad de opciones, la repercusión de las

decisiones, etc. Vincular la escuela con el entorno del alumnado, de forma integradora y dinámica, es uno de los objetivos de las pruebas.

4.1. El ítem-problema

En coherencia con el planteamiento anterior, se propone un modelo de ítem con las siguientes características:

- *El ítem-problema.* Esta opción se basa en seleccionar aspectos relevantes de la vida cotidiana, de la actualidad, del entorno –espacial, generacional, social, educativo, de los intereses...– del alumnado. La ventaja es que, en buena medida, representan virtualmente los problemas y situaciones de la vida real y, por tanto, ofrecen una evaluación acerca de las estrategias que la alumna o alumno activa y de las destrezas, recursos y actitudes que emplea para su análisis y resolución. Crea, pues, un contexto en el que aplicar lo aprendido en la escuela (contenidos, procedimientos, estrategias, etc.) y en su propio avatar personal, lo cual enlaza perfectamente con la filosofía que sustenta la evaluación diagnóstica sobre competencias básicas. Este tipo de ítem, que hemos denominado *ítem-problema*, permite además encadenar varias preguntas referidas al texto motivador que lo acompaña.

Por tanto, el texto que se utilizará en cada ítem es el elemento mediante el cual se plantea un problema. Éste puede ser de muy distinta naturaleza (conflicto personal o social, dificultad económica, dilema ético, problema ambiental, diversidad cultural, situación de convivencia, etc.); pero, en todo caso, debe acercarse al ámbito de experiencia (personal y educativa) del alumnado.

En algunos casos en que el aspecto que se evalúe sea estrictamente matemático o lingüístico, el tema del texto servirá para crear un contexto próximo al alumnado. El sentido de proximidad no debe entenderse sólo en un sentido espacial, sino también temporal y de actualidad. Así, se seleccionarán temas relevantes y apropiados a la edad del alumnado que se va a evaluar. Como se ha indicado más arriba, los temas serán variados, no estrictamente disciplinares (el objetivo no es evaluar el grado de conocimiento del currículo, puesto que la evaluación diagnóstica no es curricular), centrados en evaluar la aplicación de lo aprendido en la escuela al análisis o resolución de características de la realidad, de problemas del entorno, de aspectos relevantes de la vida actual... en distintos contextos. En cualquier caso, los ítems deben referirse a aspectos significativos, básicos, importantes.

4.2. Estructura de los ítems

La estructura habitual de un ítem será la siguiente:

1. La cabecera o presentación, al comienzo, que ofrezca una indicación de carácter general o alguna expresión que oriente sobre la finalidad de la lectura del texto o como elemento motivador.
2. El texto, que constituye la fuente informativa que se proporciona para contestar posteriormente a las preguntas. Puede ser muy variado (texto escrito, gráfico, cuadros, formularios, imágenes, esquema, tabla estadística, etc.) e incluso mixto, es decir, combinar elementos textuales de distinta naturaleza (por ejemplo, un texto escrito y un gráfico o una imagen, como aparecen frecuentemente en los libros de texto o en los medios de

comunicación escritos). El texto facilitará la información que se necesitará para contestar las preguntas; será claro y correcto desde un punto de vista lingüístico, adaptado a la edad y características del alumnado al que se dirige y, salvo excepciones que lo justifiquen, evitará alargarse con demasiadas informaciones que no se vayan a utilizar en la resolución de las preguntas.

3. *La base*. Es la cuestión específica que se plantea al alumno o alumna. Suele formularse con una frase interrogativa (“¿Cuál es el personaje que...?”, “¿Qué efecto produce el experimento que se describe en el texto?”, etc.) o mediante una instrucción u orden (“Indica el ingrediente principal de...”, “Calcula el porcentaje de materia necesaria para...”, “Subraya la idea principal”, etc.). También puede elaborarse una base mixta, en la que se proporcionan datos sobre una situación y, a continuación, se formula una pregunta, que debe resolver utilizando la información del texto general y sus propios conocimientos y destrezas.
4. *Las alternativas de respuesta*. Son las diferentes opciones entre las que el alumno o alumna debe elegir la respuesta que estime correcta. En los ítems de opción múltiple, serán cuatro. En cambio, en otras preguntas se ofrecen a veces dos respuestas alternativas (verdadero/falso, sí/no, etc.). En cualquier caso, sólo una de ellas será correcta. Las demás –los denominados distractores– serán creíbles, es decir, opciones situadas en el mismo nivel que la respuesta correcta, elaborados igualmente a partir de la información del texto y, en la medida de lo posible, reflejarán los errores habituales del alumnado de la edad que se está evaluando.

4.3. Tipos de ítems

Desde un punto de vista formal, en las pruebas de evaluación diagnóstica se proponen cuatro tipos básicos de ítems:

- ✚ *Ítem de respuesta cerrada bajo el formato de elección múltiple*. Se conoce también como *ítem cerrado*, porque cada pregunta ofrece varias opciones, con una única respuesta correcta.
- ✚ *Ítem semi-cerrado*. Son aquéllos en que la alumna o alumno debe elegir entre dos posibilidades ante una información. Por ejemplo, mediante una tabla se disponen varias afirmaciones o varias preguntas, y se debe elegir entre *sí/no*, entre *verdadero/falso*, entre *mamífero/reptil*, entre *Asia/África*, entre *natural/social*, etc. La alumna o alumno debe rodear con un círculo o marcar con una equis la respuesta que cree correcta. En todos los casos se ofrecerá un ejemplo de respuesta correcta, que guíe al alumnado en sus respuestas.
- ✚ *Ítem de construcción corta*. Se orientan a evaluar destrezas que no es posible evaluar mediante otro tipo de ítems. Su variedad es muy amplia: se le puede pedir que escriba la respuesta, que dibuje algo, que realice una operación matemática, que rellene los huecos de un esquema, que elabore un gráfico, etc.
- ✚ *Ítem abierto*. Este tipo de ítems se caracteriza por admitir una amplia gama de respuestas, las cuales, aún siendo correctas, pueden diferir según la alumna o alumno que las elabore. Para evitar el problema de la subjetividad a la hora de la calificación, este tipo de ítems van acompañados de una plantilla y criterios claros de corrección, a fin de evitar al máximo las posibles discrepancias. Los criterios de corrección gradúan las respuestas, estableciendo niveles de ejecución intermedios entre las respuestas correctas e incorrectas.

5. ¿Qué aspectos lingüísticos se han tenido en cuenta al elaborar los ítems?

La experiencia nos indica que todos los aspectos lingüísticos suelen ejercer una influencia en las evaluaciones difícil de controlar. En ocasiones, porque los ítems, aunque traten una temática científica, social o de otro tipo, se ven condicionados por la lengua utilizada. Existe la duda, en estos casos, de si se están evaluando destrezas científicas, por ejemplo, o, sobre todo, lingüísticas. Por esta razón, es recomendable extremar la atención y evitar situaciones indeseadas que disturben los resultados y su posterior interpretación.

Se ha puesto un especial cuidado a la hora de elaborar los ítems, seleccionar o adaptar los textos y redactar las opciones de respuesta, puesto que todos estos aspectos tienen un gran peso en las evaluaciones. Por tanto:

- Se ha prestado una atención especial al vocabulario y a la sintaxis, de modo que no dificulte la comprensión o lo convierta en inadecuado a la etapa en la que se utiliza. Una sintaxis recargada o un vocabulario excesivamente elevado introducen una dificultad añadida a las pruebas y originan unos resultados sesgados por el grado en que los sujetos han desarrollado su competencia para la comprensión de textos escritos, a pesar de que se esté tratando de medir otro tipo de competencias.
- Se ha adaptado el texto y la formulación de las preguntas a 4º de Educación Primaria o a 2º de la ESO, según el alumnado objeto de la evaluación. No debe incrementarse la dificultad de un ítem empleando términos inusuales para la edad.
- Se ha optado por explicaciones e instrucciones claras, directas y breves.
- Se han evitado textos largos, con pocos signos de puntuación y la “sobrecarga conceptual”, que sólo contribuyen a oscurecer el texto o la pregunta.

Un aspecto relevante en las pruebas de evaluación que se realizan en el País Vasco es la lengua de uso en los ítems (euskara y castellano). Se ha procurado un uso del euskara y del castellano en las pruebas lo más igualitario posible, de forma que el empleo de una u otra lengua no introduzca dudas sobre el grado de dificultad de un ítem y, por consiguiente, empañe la fiabilidad de los resultados.

Por estas razones, todos los ítems que se vayan a utilizar en las pruebas han pasado un control lingüístico exhaustivo antes y después de su aplicación en una prueba piloto, de manera que se garantice la calidad de las versiones en euskara y castellano, el contraste de criterios y la idoneidad a la edad del alumnado evaluado.

- A la hora de crear ítems se ha acudido a fuentes tanto en euskara como en castellano, evitando expresamente que una mayoría de los ítems proceda de una de las dos lenguas.
- La realización de las versiones en euskara y castellano garantizará lo más posible su similitud y que el grado de dificultad de una y otra no difieran sustancialmente, sin perjuicio del empleo de las características que a cada lengua le son propias. En resumen, se evitará que la alumna o alumno que emplee una de las dos lenguas se vea beneficiado o perjudicado y que, por tanto, se reste fiabilidad a la interpretación de los resultados.
- Se han evitado tanto las traducciones o versiones literales, que suelen generar mayor dificultad, porque, aun siendo correctas, no se adaptan a la edad del destinatario (es frecuente encontrar traducciones correctas, pero que son más

“cultas” o “eruditas”, que presentan frases encadenadas...), como las versiones excesivamente “libres”, que originan que una de ellas sea más difícil que la otra.

- En las versiones en euskara, se han tenido en cuenta las normas del euskara *batua*, pero se ha puesto cuidado en conseguir la comprensión del alumnado habituado a las diferencias dialectales estandarizadas.