

## Evaluaciones Externas Internacionales del Sistema Educativo



### Bloque 3. Evaluaciones externas internacionales

#### Autores

David Cervera Olivares  
Paloma González Chasco  
José M<sup>a</sup> Sánchez Echave  
INEE

## Índice

Evaluaciones externas internacionales .....	2
Aplicación y trabajo de campo .....	2
Población y muestra .....	2
Aplicación de la prueba.....	4
Instrucciones para equipo directivo, aplicador y alumnos .....	5
Ejemplos de distribución horaria .....	7
PIRLS y TIMSS .....	7
PISA.....	7
EECL.....	8
Funciones del aplicador y Situaciones especiales .....	9
Instrucciones para el desarrollo de la aplicación .....	10
Importante.....	10
Corrección de la prueba (codificación).....	11
Ejemplo de TIMSS ciencias .....	11
Nutrientes.....	11
Ejemplo de TIMSS ciencias: criterios de corrección.....	12
Nota .....	14
Ejemplo de PISA.....	14
De mal gusto.....	14
Ejemplo de PISA: criterios de corrección .....	15
¿Te atreves a codificar las respuestas?.....	17
Descarga los criterios de corrección.....	17
Corrección del ejemplo "De mal gusto" .....	17
EECL.....	20
¿Quién escribió "El Quijote"?.....	22

## Evaluaciones externas internacionales

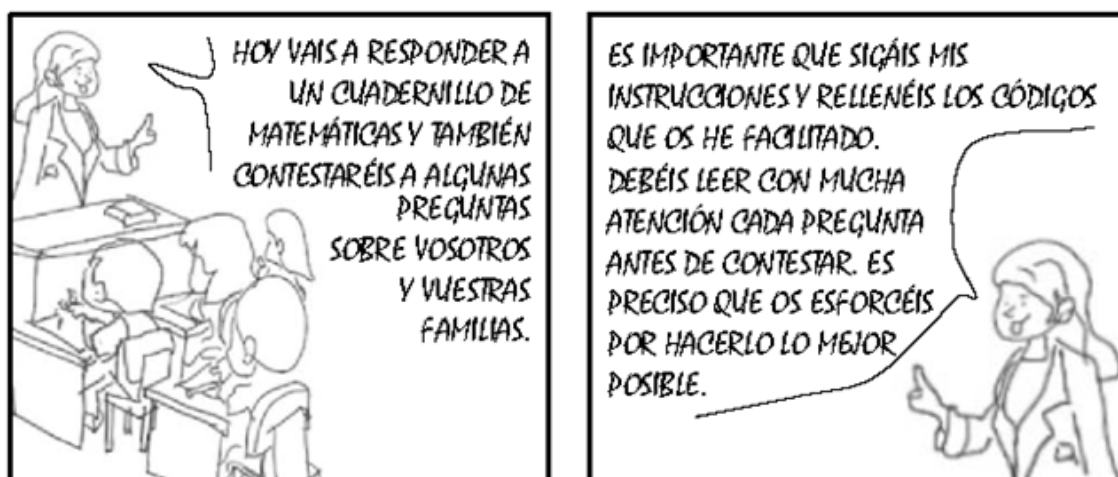
### Aplicación y trabajo de campo

#### Aplicación y trabajo de campo

- Población y muestra.
- Aplicación de la prueba.
- Corrección de la prueba.

El **proceso de aplicación** es el momento en el que a través de los alumnos se consigue la información que permitirá un análisis posterior. Es importante destacar, aunque ya lo hemos mencionado anteriormente, que las evaluaciones externas internacionales no pretenden evaluar o calificar al alumno, sino que buscan analizar los sistemas educativos y detectar vías de mejora a partir de la observación de buenas prácticas que puedan ser replicadas.

Para ello se emplean las respuestas de los alumnos a las **pruebas cognitivas** y a los **cuestionarios de contexto**, que permiten extrapolar y analizar los resultados en el rendimiento a través de diversos indicadores como género, contexto social, situación económica, etc., pudiendo a través del análisis detectar fortalezas y debilidades. Por tanto, una vez configurada la prueba, dividida en sus diferentes cuadernillos o módulos informáticos y preparados los cuestionarios de contexto, se hace **imprescindible ir al aula** y contar con la colaboración del alumnado y su participación con implicación en la realización de los ejercicios y las distintas partes que configuran estas evaluaciones.



Si no ves correctamente la imagen anterior, puedes descargarla este [PDF](#).

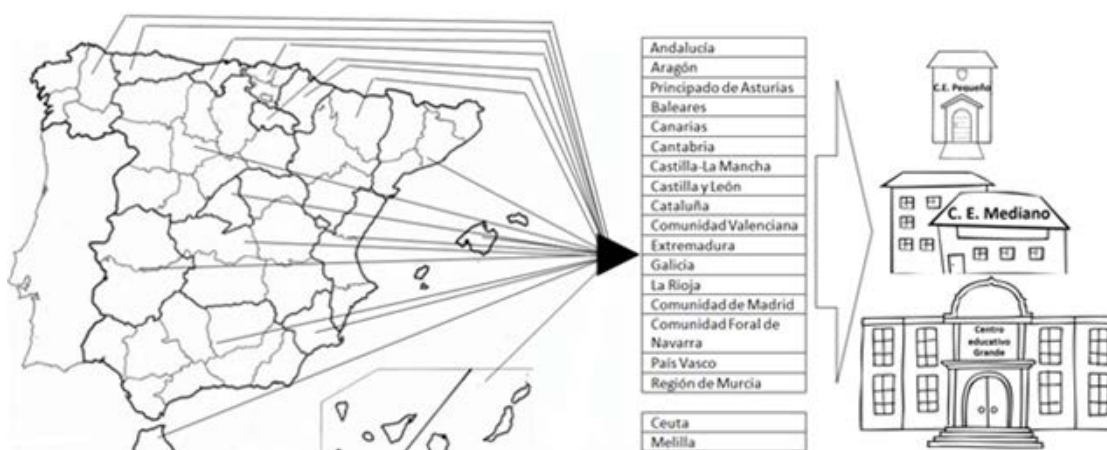
### Población y muestra

El marco teórico de la evaluación determina las características del alumnado que realiza las pruebas. Según cada evaluación internacional podemos ver los siguientes ejemplos:

- **PIRLS y TIMSS:** grupos completos de alumnos de 4º de Educación Primaria, seleccionados aleatoriamente. Un grupo en caso de centros pequeños y dos de los grupos si el centro es grande. Consideramos que un centro es pequeño, mediano o grande en función de la línea(número de clases de un mismo nivel) que tienen para el curso en el que se realiza la evaluación..
- **PISA:** 35 alumnos seleccionados al azar de entre todos los del centro que cumplen 16 años en ese curso escolar. (En el caso de España alumnos de 4º, 3º o 2º de ESO, siempre que cumplan la condición de 15 años para cumplir 16 en ese curso escolar).
- **EECL:** 25 alumnos del conjunto de alumnado de 4º de Educación Secundaria Obligatoria con al menos un año en el estudio de la lengua evaluada.

No todos los centros forman parte de la muestra. Los **centros son seleccionados aleatoriamente** de forma que quede representada la realidad que se quiere analizar. En España, en general, se consideran los siguientes **estratos** (divisiones o clasificación en grupos de los que se extraen porcentajes representativos): las 17 **comunidades autónomas** y las ciudades autónomas de Ceuta y Melilla; de todos ellos deben existir alumnos. Además en cada Comunidad Autónoma se organizan los centros según **titularidad**, públicos, concertados y privados. Y dentro de estos tres grupos se subdividen en centros **pequeños, medianos y grandes**. Una vez asignados todos los centros del territorio español a cada grupo según esta clasificación, se seleccionan aleatoriamente de estos grupos de forma que el porcentaje de representación de cada uno de los grupos se corresponda con la realidad a analizar.

Como algunas **CCAA amplían muestra** en algunas evaluaciones para poder obtener datos representativos, además se puede hablar de otros estratos como las provincias, direcciones de área territorial (Madrid), servicios periféricos (Castilla-La Mancha), delegaciones provinciales, ..., según la comunidad.



Los **centros seleccionados** que forman parte de la muestra son informados a través de sus respectivas administraciones educativas de la participación en una u otra prueba al haber sido seleccionados como centro adjudicado. Esta información se facilita en los dos o tres meses previos a la realización de la prueba. No todos los centros realizan la prueba el mismo día dado que existe una persona (**aplicador**) que irá al centro a realizar la evaluación. Con suficiente antelación se acordará junto con el equipo directivo del centro el momento más adecuado para

hacerla dentro del periodo establecido por el organismo internacional correspondiente (OCDE, IEA, UE).

**No existe riesgo de difusión** de las preguntas dado que los aplicadores entregan y recogen todo el material y el profesor que acompaña, en ningún momento tiene acceso a fotocopiar o digitalizar los cuadernillos de trabajo de los alumnos. Hay que señalar que existen distintos cuadernillos y preguntas diversas dentro de cada uno de ellos, lo que dificulta aún más el proceso de posible copia en caso de que se produzca alguna incidencia.

Estas **pruebas son confidenciales** y no pueden ser fotocopiadas ni difundidas bajo ningún concepto, por lo que el aplicador se responsabiliza de su custodia y firma un compromiso de confidencialidad. Por otro lado, los centros participantes generalmente no tienen información directa de qué otros centros forman parte de la muestra. De esta forma, aunque la prueba se realiza en días distintos, la posibilidad de filtración de preguntas es realmente difícil.

## Aplicación de la prueba

Conocida la muestra de centros y de alumnos dentro de cada centro llega el momento de ir al colegio o instituto y dar instrucciones a los alumnos, repartir los cuadernillos, etc. Una pregunta habitual que el alumnado realiza cuando se les reparten los cuestionarios es: "**¿dónde pongo mi nombre en el examen?**" La respuesta es "**En ningún sitio**". Para **garantizar el anonimato** de los alumnos los centros asignarán un código a cada uno que será una identificación que únicamente el responsable del centro puede conocer en relación con el nombre del alumno. Ningún cuadernillo sale del centro con un nombre que lo identifique y la conexión código-alumno, solo la conoce el responsable del centro educativo.

Como se ha señalado, no todos los alumnos realizan las pruebas, ni estas son las mismas para aquellos que participan, por lo que **no es relevante devolver una nota al alumno**. Sería injusto, dado que el grado de dificultad de las distintas partes (cuadernillos) que configuran la prueba total es variable. Además, los datos no tienen suficiente significatividad para devolver resultados objetivos al centro, dado que solo una pequeña parte de su alumnado ha participado. En cualquier caso es interesante que el centro conozca estos resultados, una vez finalizado el proceso. Existe una importante información subjetiva que se puede analizar en los departamentos conociendo el marco teórico en el que está fundamentada la prueba y conociendo los ítems liberados. Hay que destacar que las comparaciones directas no se deben realizar dado el elevado error estadístico que tendríamos.

Dada la complejidad del proceso de pilotaje en estos estudios debida a factores como el número de países participantes, las fechas en las que cada país participa, la recogida de toda la documentación, la corrección y codificación, etc., hace que las bases de datos tarden varios meses en estar preparadas para el análisis y que los **resultados** se proporcionen aproximadamente **un año más tarde**, otro factor que contribuye a no devolver datos individualizados al alumno pues en muchos casos ya ni se encontrarán en el mismo.

Algunas evaluaciones, como es el caso de **PIRLS y TIMSS**, incluyen **cuadernillos especiales** que realizan **solo algunos alumnos del grupo** y que se emplean para hacer

análisis secundarios y comparativos de consistencia dentro de la misma prueba. Por ejemplo en PIRLS, todos los alumnos se enfrentan a los textos en blanco y negro y con las preguntas situadas a continuación de cada uno de los textos. En cambio algunos alumnos disponen de un cuadernillo a todo color con todos los textos seguidos, como si de un libro se tratase y otro cuadernillo en el que se encuentran las preguntas y en el que tiene que contestar.

## Instrucciones para equipo directivo, aplicador y alumnos

Dependiendo de la prueba el proceso de aplicación varía, aunque en casi todas ellas se emplean procesos estandarizados de forma que se **garantice la homogeneidad de la aplicación** en todos los centros.

Podríamos hablar de instrucciones al menos en tres niveles aunque son similares con adaptaciones puntuales según el rol de cada participante:

- Instrucciones para el equipo directivo.
- Instrucciones para el aplicador.
- Instrucciones para el alumno.

A continuación se muestran algunas de las normas generales. Para conocer en detalle las normas específicas de cada prueba habría que consultar los manuales del aplicador de cada evaluación.

Los **equipos directivos** deben estar familiarizados con el proceso de la prueba, conocer los **tiempos** para adaptar la / las jornadas lo mejor posible a la propia jornada habitual del centro, por ejemplo haciendo coincidir la hora del recreo, la hora de la comida, etc. Los tiempos dedicados a la realización de la **prueba cognitiva** no se pueden variar, pero se puede modificar el comienzo para hacerlo coincidir, por ejemplo, con la hora de cambio de clase. Los tiempos de **descanso** tampoco se deben modificar pero en esto se puede ser flexible en unos pocos minutos si facilita la adaptación a la organización del centro. Por otro lado, los **cuestionarios de contexto** deben cumplimentarse de la mejor forma posible. Si esto significa dar algo más de tiempo a los alumnos para la recopilación de esta información podemos alargar los tiempos.

El director del centro es el responsable de los alumnos y será quien determine quién estará en el aula durante la aplicación: él mismo, un jefe de estudios, el tutor del grupo, etc.

El **aplicador** es el responsable de la prueba pero no de los alumnos, y no debe estar solo en ningún momento del proceso. Si durante la aplicación se produjera algún incidente por el que fuera necesario que alguien abandonara el aula, debe estar el docente asignado como responsable del grupo para que la incidencia altere lo menos posible al proceso de la aplicación.

El aplicador se pondrá en contacto con el centro educativo y contrastará que está informado de la realización de la prueba, le explicará las fases al responsable para comprobar que ambos

conocen los procesos de funcionamiento y acordarán una hora previa al comienzo de la evaluación para verificar la disposición individual de mesas, el espacio y otras características del centro. Es responsabilidad del aplicador llevar y recoger los materiales, y firmar y solicitar la firma de algún miembro del equipo directivo de determinados procesos realizados, como la asistencia de los alumnos o la no asistencia, características especiales de alguno de ellos, ausencias justificadas por visitas de médico, así como otras incidencias inesperadas, etc.

El aplicador únicamente conoce el código del alumno por lo que es el **responsable del centro** el que debe proporcionar e **identificar a cada alumno al comienzo de la prueba**, aunque el aplicador realizará la verificación del proceso. Por ejemplo, el listado con nombres y apellidos de los alumnos con los códigos solo los tiene el centro, pero durante la aplicación los alumnos se sentarán en orden alfabético. Incluso si por algún motivo un alumno falta, se dejará el hueco y se pegarán etiquetas con los códigos de los alumnos en cada mesa. El aplicador tendrá a su disposición este listado que vincula código con nombre y puede solicitar al alumno su identificación, pero no podrá quedarse con este documento.

El **aplicador** es el encargado de **transmitir las instrucciones** de funcionamiento de la jornada **a los alumnos**, de responder sus preguntas, de orientarles, animarles y guiar en definitiva todo el proceso. El **docente** que acompañe al grupo debe encargarse del correcto funcionamiento y el **orden del aula**, no interfiriendo en las labores del aplicador ni en el reparto de materiales, ni en su recogida y no debe contestar ninguna pregunta de los cuadernillos cognitivos. En cambio toda **ayuda** será poca para la correcta **interpretación** de las preguntas de los **cuadernillos de contexto** y su ayuda en este proceso puede ser de gran utilidad, sobre todo si es tutor o profesor del grupo. Un procedimiento que en ocasiones se emplea para cumplimentar los contextos es que el aplicador guíe todo el proceso de respuestas en el cuadernillo, leyendo una a una y explicando las dificultades, que seguro surgen. De esta forma garantizamos que todos los alumnos entienden y contestan lo más verídicamente posible a las cuestiones sobre su contexto. También puede dejar que contesten de manera individual pero aclarando todas las dudas y solicitando ayuda, si fuera necesario del docente, para agilizar el proceso a la hora de que los alumnos cumplimenten el cuestionario.

Las **instrucciones para los alumnos** se recogen en la primera o primeras páginas de los cuadernillos y se refieren a cómo contestar, alguna pregunta tipo a modo de ejemplo, duración, cómo tachar y corregir una pregunta en la que se han confundido, etc. Las instrucciones son para los alumnos pero el aplicador debe conocerlas y debe informar y resolver las dudas derivadas de las mismas antes de comenzar el desarrollo de la prueba.

Como se ha señalado en varias ocasiones, los alumnos seleccionados para formar parte de la muestra contestan además de las pruebas, un cuestionario de contexto y en ocasiones estos se complementan con **cuestionario de centro, profesor y familias** en función de la prueba. En el estudio PIRLS y TIMSS, el cuestionario para las familias fue entregado a los niños y niñas al terminar el primer día de la prueba, con el ruego de que lo devolvieran cumplimentado a su profesor tutor en los próximos días. Muchos de ellos fueron recogidos en el segundo día de prueba, asegurándose así una alta tasa de participación; otros se recogieron con posterioridad gracias a la colaboración de los equipos directivos.

Existen documentos que deben firmar tanto aplicador como responsable del equipo directivo como garantía de calidad del proceso. Además en un porcentaje no menor al 10% se realizan **controles de calidad**, por lo que es posible que acudan al centro miembros de la organización internacional responsable, de la Administración central o de la Administración autonómica, que observarán todo el proceso de aplicación.

## Ejemplos de distribución horaria

### PIRLS y TIMSS

La prueba de 2011 se aplicó en dos días. Para neutralizar el posible efecto del cansancio en los resultados, la mitad de los centros hicieron el primer día la prueba de PIRLS y la otra mitad, la de TIMSS. La organización y el horario de las pruebas podemos verlo en la siguiente tabla:

<b>ACTIVIDAD</b>	<b>Primer día</b>	<b>Segundo día</b>
<b>Preparación de los alumnos</b> (instrucciones)	10 min	5 min
<b>Primera parte de la prueba</b>	40 min PIRLS / 36 min TIMSS	36 min TIMSS / 40 min PIRLS
<b>Descanso</b>	15 min	15 min
<b>Preparación</b>	5 min	5 min
<b>Segunda parte de la prueba</b>	40 min PIRLS / 36 min PIRLS	36 min TIMSS / 40 min PIRLS
<b>Cuestionario del alumno</b>	10 - 30 min	

### PISA

<b>ACTIVIDAD</b>	<b>Tiempo</b>



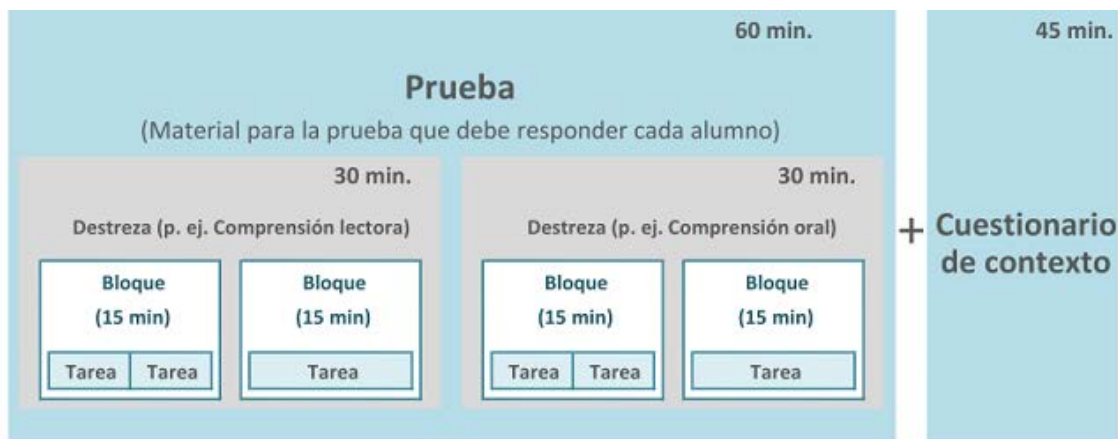
<b>Preparación de los alumnos (instrucciones)</b>	10 minutos
<b>Tiempo de la Prueba</b>	55 minutos
<b>Descanso</b>	10 minutos
<b>Tiempo de la Prueba</b>	55 minutos
<b>Descanso</b>	30 minutos
<b>Cuestionario de contexto</b>	30 minutos

En los años 2009 y 2012, 14 de los 35 alumnos, seleccionados al azar, también realizaban una prueba mediante ordenador.

<b>OPCIÓN COMPUTER BASED ASSESSMENT (prueba por ordenador)</b>	<b>Tiempos</b>
<b>Preparación</b>	20 minutos
<b>Prueba</b>	40 minutos

## EECL

A continuación podemos ver un esquema donde se representa la distribución temporal de la prueba para los bloques de contenido y el cuestionario de contexto.



Si no ves correctamente la anterior imagen, descárgate este [PDF](#).

## Funciones del aplicador y Situaciones especiales

Las **funciones** que debe cumplir el **aplicador** son las siguientes:

- Representación del organismo internacional para la realización de la prueba.
- Garantizar el desarrollo de la prueba en un espacio adecuado de trabajo.
- Velar por el cumplimiento de las instrucciones.
- Crear un clima adecuado de atención e interés.
- Transmitir con precisión instrucciones generales y específicas.
- Resolver las dudas de manera global y a ser posible antes de comenzar la prueba.
- Indicar tiempo, hora de inicio y hora de fin.
- Observar como rellenan los alumnos los datos.
- Hacer un control horario preciso de cada parte de la prueba.
- Respetar el orden de las pruebas y sus partes.
- Ayudar en el cuestionario de contexto.
- Motivar a los alumnos para que contesten con interés.
- Mantener el silencio y el orden.

Cuando en un centro participe un solo **grupo**, de entre los que tiene, la **selección** se realizará **al azar** por el organismo internacional responsable, al igual que cuando la selección se realiza de manera individual sobre los **alumnos**, como ocurre en PISA.

Algunas posibles **situaciones especiales** y **procedimientos aplicables** durante el proceso de aplicación:

- Si un **alumno llega tarde**, no debe incorporarse a esa parte de la prueba, haciéndolo, sin embargo, a las siguientes., registrando la incidencia.
- En caso de que un **alumno** tuviera que **abandonar** la prueba se debe anotar en la hoja de seguimiento la razón del abandono.

- Los alumnos con **necesidades educativas especiales** o de **compensación educativa** realizarán la prueba si así lo recoge el marco teórico y las instrucciones de la misma, quedando esta información recogida en el formulario de seguimiento de la sesión (estos alumnos no intervienen en los cálculos finales de rendimiento).

En el **formulario de seguimiento de la sesión** se recogerán las incidencias anteriores y otras referidas a los alumnos.

## Instrucciones para el desarrollo de la aplicación

Algunas **instrucciones comunes** al desarrollo de aplicación en distintas pruebas:

- Los **alumnos** serán colocados por **orden de lista**.
- Cuando los alumnos estén colocados en el aula, el aplicador dará las **instrucciones generales**, tratando de crear un clima adecuado de atención e interés, con aproximación al guión siguiente:
  - Hoy vais a responder a una serie de cuadernillos (de matemáticas, comprensión lectora y lengua) y también contestaréis a algunas preguntas sobre vosotros y vuestras familias. Debéis responder en ...
  - Es importante que sigáis mis instrucciones en todo momento y rellenéis los códigos que os diga en la hoja de respuestas y en el cuestionario de contexto.
  - Debéis leer con mucha atención cada pregunta para entender con claridad lo que se os pide que hagáis.
  - Aunque los resultados de estas pruebas no van a influir en las calificaciones de final de curso, es preciso que os esforcéis por hacerlo lo mejor posible.
  - Deberéis realizar los ejercicios en silencio; si tenéis alguna duda, levantad la mano y esperad a que se os atienda.
  - Si no entendéis o no sabéis responder a una o más preguntas, después de haberlas leído atentamente, debéis pasar a las siguientes. Al final, si os queda tiempo, podréis intentar de nuevo responderlas.

### Importante

**EL APLICADOR DEBE PONER ESPECIAL INTERÉS EN QUE LOS ALUMNOS MARQUEN CORRECTAMENTE LOS CÓDIGOS EN LOS CUADERNILLOS COGNITIVOS, EN LOS CUESTIONARIOS DE CONTEXTO, EN LAS HOJAS DE RESPUESTA, SI SE REALIZA POR ORDENADOR QUE LA INCLUSIÓN DE LOS CÓDIGOS SEA CORRECTA, ETC.**

Para el **día de la aplicación** se recomienda:

- Concertar con el aplicador la llegada al centro con antelación suficiente para solucionar posibles imprevistos (20 min).
- Solicitar la acreditación como aplicador de la prueba.

- Verificar junto con el aplicador que se dispone de todo el material necesario para la aplicación de la prueba y, en caso de incidencia, registrarla en la hoja de seguimiento y expresar la solución adoptada, contactando en caso necesario con la Administración responsable.
- Repasar conjuntamente el proceso de aplicación y presentar a las personas que estarán en el aula durante la prueba. El aplicador no debe de estar solo en el aula en ningún momento.
- Comprobar que el aula cumple las condiciones necesarias para la prueba junto con el aplicador.
- Aunque el aplicador controlará los tiempos con exactitud es conveniente que el aula disponga de reloj.
- Realizar las adaptaciones de los horarios del centro con los de la prueba. Se debe respetar en todo momento los tiempos de cada parte y ajustando los descansos a periodos como el recreo o la comida.

## Corrección de la prueba (codificación)

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. Por ejemplo, penalizar **por faltas de ortografía o por una mala expresión gramatical**. En las pruebas internacionales es difícil la comparación entre lenguas por lo que no se debe añadir un elemento que distorsione o rompa la homogeneidad, al participar diversos países y emplear distintas lenguas para expresar las respuestas. 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.

## Ejemplo de TIMSS ciencias

Observa en el siguiente ejemplo de TIMSS cómo son las reglas de codificación.

### Nutrientes

Gregorio le cuenta a su amiga Sofía que él puede obtener de la fruta todos los nutrientes que necesita para estar sano. Sofía piensa que Gregorio tiene que tomar, además, otros tipos de alimentos.

¿Quién está en lo cierto?

(Marca una casilla.)

o Gregorio

o Sofía

Explica tu respuesta

---



---



---



---

Regla de codificación

**NOTA:** Se consideran correctas tanto aquellas respuestas que mencionen alimentos concretos (o grupos de alimentos) necesarios para seguir una dieta equilibrada, como las respuestas que hagan referencia a nutrientes concretos que no se obtienen con una dieta basada únicamente en la fruta (código 11). Las respuestas que solo repitan la información dada en el enunciado (diciendo que se necesita comer otros tipos de alimentos, o que comer solo fruta no es saludable, etc.) se considerarán incorrectas. Al corregir la respuesta debe primar la explicación. Si esta es correcta, la respuesta debe considerarse correcta aunque el alumno haya marcado GREGORIO (Código 12), lo que indicaría cierta confusión acerca de la postura de Gregorio y de Sofía.

### Ejemplo de TIMSS ciencias: criterios de corrección

Código de corrección	Criterios de corrección en las respuestas
<b>Respuesta correcta</b>	
<b>10</b>	<p>SOFÍA, junto con una explicación que indique otros alimentos (o grupos de alimentos) que necesita comer Gregorio.</p> <p>Ejemplos: Verduras, carne, pescado, huevos, judías, arroz, cereales, productos lácteos, etc.</p>
<b>11</b>	<p>SOFÍA, junto con una explicación que indique nutrientes que faltan en una dieta basada únicamente en la fruta.</p> <p>Ejemplos:</p>

	<p>Hay que comer diferentes alimentos para obtener todas las vitaminas.</p> <p>Se necesitan algunas proteínas para mantenerse sano.</p>
<b>12</b>	<p>GREGORIO, junto con una explicación correcta.</p> <p>Ejemplos:</p> <p>Tiene que comer alimentos de los cuatro grupos para que la dieta sea sana.</p>
<b>19</b>	Otras respuestas correctas.
<b>Respuesta incorrecta</b>	
<b>70</b>	<p>SOFÍA, sin explicación o con una explicación incorrecta [se incluyen respuestas que solo repitan la información del enunciado, o que digan algo verdadero que no responda a la pregunta.]</p> <p>Ejemplos:</p> <p>Para tener buena salud también son necesarios otros tipos de alimentos.</p> <p>Es malo comer demasiada fruta.</p> <p>Creo que Sofía tiene razón.</p> <p>Necesita otros alimentos para obtener nutrientes.</p>
<b>71</b>	<p>GREGORIO, sin explicación o con una explicación incorrecta [se incluyen respuestas que solo repitan la información del enunciado, o que digan algo verdadero que no responda a la pregunta.]</p> <p>Ejemplos:</p> <p>La fruta contiene muchas vitaminas.</p>
<b>79</b>	Otras respuestas incorrectas (incluidas respuestas tachadas/borradas, marcas fuera de su sitio, respuestas ilegibles o inapropiadas).

<b>Sin respuesta</b>	
<b>99</b>	En blanco

### Nota

En realidad, observa que 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**. Es importante destacar que 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. En el caso de PIRLS y TIMSS, además del proceso de corrección en las preguntas abiertas, se realiza al 25 % un proceso de **doble corrección** de las respuestas totales asegurando que el grado de acuerdo entre correctores supera la tasa requerida. Por otro lado algunos correctores con dominio de inglés realizan una corrección transversal a un porcentaje de respuestas en idioma inglés. Con este doble ejercicio se garantiza la fiabilidad de los criterios de corrección aplicados en los países participantes.

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.

Puedes consultar los estímulos de **PIRLS y TIMSS** que se han liberado del último estudio realizado en 2011. Presta especial atención a las guías de codificación. Para ello consulta el anexo del capítulo 1, ítems liberados del volumen “PIRLS - TIMSS 2011. Estudio Internacional de progreso en comprensión lectora, matemáticas y ciencias. IEA. Volumen I: INFORME ESPAÑOL” (págs. 141 - 180). Se puede descargar en el siguiente [enlace](#) o en este otro [alternativo](#).

### Ejemplo de PISA

El texto que aparece a continuación es uno de los textos continuos en prosa más breves del conjunto de los liberados de PISA. Es una carta titulada “De mal gusto” que apareció en un periódico y ha sido clasificada como texto argumentativo porque el autor intenta convencer al lector de su propio punto de vista. Procede de un contexto público.

#### De mal gusto

Esta carta apareció en un periódico australiano en 1997. Remítete a ella para responder a las siguientes preguntas.

¿Sabías que en 1966 los australianos gastaron casi la misma cantidad en chocolate que gastó su gobierno en ayuda exterior?

¿Hay algo que no funciona en nuestra escala de valores?

¿Qué vas a hacer al respecto?

Sí, tú.

Arnold Jago,

Mildura

Fuente: The Age, 1 de abril de 1997

### Pregunta 1

El propósito de la carta de Arnold Jago es provocar

- A. Culpabilidad
- B. Diversión
- C. Temor
- D. Satisfacción

### Pregunta 2

¿Qué tipo de respuesta o de acción crees tú que quiere provocar Arnold Jago con esta carta?

---

---

---

---

## Ejemplo de PISA: criterios de corrección

<b>Código de corrección</b>	<b>Criterios de corrección de las respuestas</b>
-----------------------------	--



<b>Respuesta correcta pregunta 1</b>	
<b>1</b>	Opción A: Culpabilidad
<b>Respuesta incorrecta pregunta 1</b>	
<b>0</b>	Otras
<b>El alumno no responde</b>	
<b>9</b>	Sin respuesta
<b>Respuestas correctas pregunta 2</b>	
<b>1</b>	<p>Respuestas del tipo a o b</p> <p>a. Respuestas que aportan una afirmación o frase que indique que el gobierno o los particulares deberían gastar más en ayuda exterior, por ejemplo:</p> <ul style="list-style-type: none"> <li>• “Gente que dé dinero para ayuda exterior”.</li> <li>• “Dar dinero para obras de caridad”.</li> <li>• “La gente debería gastar menos en chocolate y más en los pobres”.</li> </ul> <p>b. Respuestas que aportan una afirmación o frase que indique que el gobierno o los particulares deberían cambiar su sistema de valores o su visión del problema:</p> <ul style="list-style-type: none"> <li>• “Cambiar nuestros valores”.</li> <li>• “Querría que la gente se preocupara más de cómo gastamos los recursos”.</li> </ul>
<b>Respuestas incorrectas pregunta 2</b>	
<b>0</b>	<p>Respuestas del tipo c, d o e:</p> <p>c. Respuestas que indiquen la intención del autor para hacer que el</p>

	<p>lector se sienta culpable</p> <ul style="list-style-type: none"> <li>• “Sentirse culpable o avergonzado”.</li> </ul> <p>d. Respuestas que aportan una afirmación o frase que indique que se debería gastar menos en chocolate o ser menos goloso, por ejemplo:</p> <ul style="list-style-type: none"> <li>• “No comprar más chocolate”.</li> <li>• “Dejar de comer comida basura”.</li> </ul> <p>e. Otras, vagas, inapropiadas o irrelevantes, por ejemplo</p> <ul style="list-style-type: none"> <li>• “Gastar más en obras de caridad”.</li> <li>• “Le gustaría que despidieran a los del gobierno”.</li> <li>• “Le gustaría que la gente dijese: voy a dar todo mi dinero para obras de caridad”.</li> <li>• “No estoy de acuerdo con Arnold Jago”.</li> <li>• “Estoy de acuerdo con él”.</li> </ul>
<b>El alumno no responde</b>	
<b>9</b>	Sin respuesta

En el caso del programa internacional PISA, la **corrección** de cada ejercicio se realiza por **equipos de 4 personas de manera rotatoria**, siendo la nota válida internacionalmente la del cuarto codificador. Al ir rotando, todos codifican en algún momento como cuarto y este código es el que se marca en el cuadernillo; el resto se recogen en un estadiillo que posteriormente se somete a un control de calidad. Se selecciona aproximadamente el 25% de los ejercicios al azar para comprobar que los distintos codificadores no están introduciendo ningún sesgo.

## ¿Te atreves a codificar las respuestas?

### Descarga los criterios de corrección

Bájate este [PDF](#) con los criterios de corrección del ítem 2 del estímulo PISA "De mal gusto" para que lo puedas consultar cuando hagas el siguiente test.

### Corrección del ejemplo "De mal gusto"

**Gente que gasta menos dinero en chocolate y más en los enfermos del extranjero.**



1



0



9

**Que la gente no gaste todo el dinero en chocolate en vez de en el extranjero.**



1



0



9

**Un aumento en el gasto del gobierno en ayuda exterior para ayudar a los pobres.**



1



0



9

**En vez de comprar y comer chocolate la gente debería dar para una buena causa y no ser tan indulgente consigo misma.**



1



0



9

**Gente conmovida para pensar más en ayudar a los demás que en dedicarse a gustos personales.**



1



0



9

**Que la gente se dé cuenta de que los pobres necesitan nuestra ayuda para que hagan algo sobre este asunto.**



1



0



9

**Creo que le gustarían cartas con opiniones escritas y lo que deberían hacer para ayudar este problema.**



1

●

0

●

9

Puede querer ver las sugerencias de la gente sobre cómo recaudar fondos para el uso del chocolate en el extranjero.

●

1

●

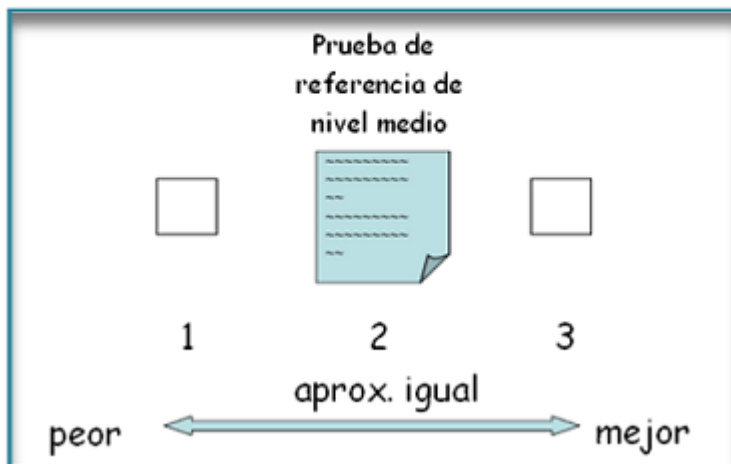
0

●

9

## EECL

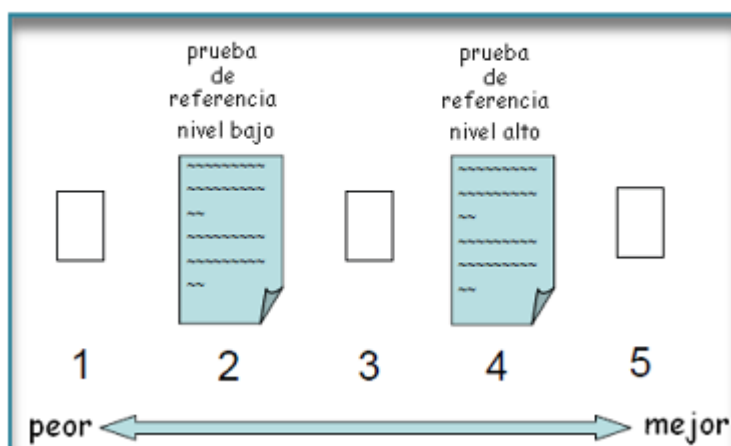
En el estudio de Competencia Lingüística en idioma extranjero **EECL**, se evaluaron tres destrezas: comprensión lectora, comprensión oral y expresión escrita. Las dos primeras se codifican de manera similar a los ejemplos que hemos visto. En cambio, para poder **codificar la expresión escrita** y evitar la subjetividad que un corrector puede incorporar al ser más duro en la codificación, se recurrió a **testigos de comparación**. Para corregir los ejercicios de los alumnos de niveles bajos se empleó un único testigo de forma que el corrector tenía que decidir si la producción del alumno (la redacción que ha escrito como respuesta) era peor, más o menos igual o mejor que la de comparación. De esta forma un corrector duro siempre tendría presente el documento empleado como testigo de comparación; tal vez este ejemplo le podía parecer muy malo pero la estrategia es decidir si la producción que está codificando es aún peor, más o menos igual o mejor. De la misma forma un profesor que no diera gran importancia a ciertos errores por considerar que el texto se entiende, puede tener la impresión de que la producción del alumno está bien, pero tendrá que emitir un juicio sobre si está mejor, más o menos igual o peor que el testigo.

**Figura 2.1. Corrección de las tareas de nivel A1-A2**


Si no ves correctamente la imagen anterior, descárgate este [PDF](#).

Es importante destacar la importancia de la formación en este tipo de tareas para que la corrección sea lo más homogénea posible, no solo a nivel nacional sino también internacional.

Cuando el nivel de producción es B1-B2, empleando expresiones gramaticales de mayor dificultad, vocabulario más complejo, etc., se hace necesario el empleo de **dos testigos** (uno de nivel bajo y otro alto). Esto es, a la hora de codificar un ejercicio si nos parece mejor que el testigo de nivel bajo, tendremos que compararlo nuevamente con el siguiente para comprobar si es peor, igual o mejor.

**Figura 2.2. Corrección de las tareas de nivel B1-B2**


Si no ves correctamente la imagen anterior, descárgate este [PDF](#).

En su origen estos testigos de comparación fueron producciones artificiales que contenían errores pero que sirvieron para los procesos de pilotaje. En el caso del estudio principal los testigos de comparación fueron producciones reales de alumnos seleccionadas de entre las que los alumnos realizaron durante el proceso de pilotaje.

Puedes consultar todos los ítems liberados y los audios de los mismos en la siguiente dirección <http://www.mecd.gob.es/inee/Recursos.html>.

No solo es importante conocer lo que se pregunta y cómo se pregunta, sino también conocer lo que se busca a la hora de la respuesta, pues la codificación depende en gran medida de esto.

## ¿Quién escribió "El Quijote"?



Si no ves correctamente la imagen anterior, puedes descargar este [PDF](#).

Pongamos como **ejemplo final** algo que puede ayudarnos a diferenciar entre **cómo corregir** un examen de **aula** por el profesor y una **evaluación internacional**.

Supongamos una pregunta.

### ¿Quién escribió "El Quijote"?

Estas son algunas de las respuestas que nos podríamos encontrar

1. Cervantes
2. Miguel de Cervantes
3. Miguel de Cervantes Saavedra
4. El Ingenioso hidalgo don Quijote de la Mancha fue escrito por el ilustre escritor Miguel de Cervantes Saavedra.
5. Servantes
6. Zervantes
7. Cerbantes

8. Góngora
9. Quevedo
10. Y yo que sé. Pregunta a otro. ¿Para que sirve esto?

¿Qué codificación pondrías?

A pesar de que las **respuestas 1, 2, 3 y 4** difieren en la cantidad de información que el alumno ofrece, no podríamos considerar que estén mejor o peor unas que otras dado que la codificación es Bien (1), Mal (0) o sin respuesta (9). Por tanto, la codificación para todas ellas sería (1).

Si nos fijamos en las **respuestas 5, 6 y 7**, la codificación en una prueba de estas características tiene que ser acordada antes de su aplicación y, si consideramos que lo importante es que el alumno sepa que Cervantes fue quien escribió el Quijote, deberían ser codificadas como correctas (1).

Por supuesto si esta pregunta la hacemos en un examen de clase, el profesor podrá pedir al alumno que sepa escribir correctamente el nombre y considerarlo como un error, pero el valor ortográfico y gramatical no es comparable entre idiomas como el castellano, el inglés, el francés, el chino, el ruso... todos ellos participantes en este tipo de pruebas.

En cambio las **respuestas 8 y 9** son incorrectas y por tanto las codificaríamos con un (0). También se codificaría con un cero **la respuesta 10**, aunque sea un sin sentido o una respuesta no vinculada a la pregunta, el alumno no la ha dejado en blanco, por lo que el codificador la identificará con un (0).



INEE - Formación en Red del INTEF

Licencia [Creative Commons Atribución-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-sa/4.0/).