

Principios y prácticas de evaluación del IB: guía sobre la evaluación para profesores y coordinadores

Introducción

Esta publicación está pensada a modo de guía resumida para profesores y coordinadores sobre los métodos que utiliza el IB para garantizar que las evaluaciones que a las que se someten los alumnos sean significativas y justas.

Esto implica:

- Mantener la coherencia en las correcciones
- Mejorar la calidad de la corrección
- Principios de consulta sobre los resultados

Para obtener información más detallada acerca de la evaluación en el IB, se puede visitar el sitio web del IB: www.ibo.org.

Pasos en el ciclo de vida de una evaluación

El IB utiliza varios procesos y controles de calidad para garantizar que las puntuaciones y las calificaciones que reciben los alumnos en todas las asignaturas y componentes sean fiables, justas y correctas. Se considera que una evaluación es *fiable* cuando el alumno obtendría el mismo resultado si repitiera la evaluación en distintas ocasiones y si la evaluación la corrigieran distintos examinadores.

En cada fase del ciclo de evaluación, el IB dedica un esfuerzo considerable para garantizar que las evaluaciones miden lo que deben medir (es decir, que son válidas) y que se aplican estándares de corrección precisos.

El ciclo de evaluación es un proceso continuo en el que cada paso viene formado por el anterior y sirve de base para el siguiente.

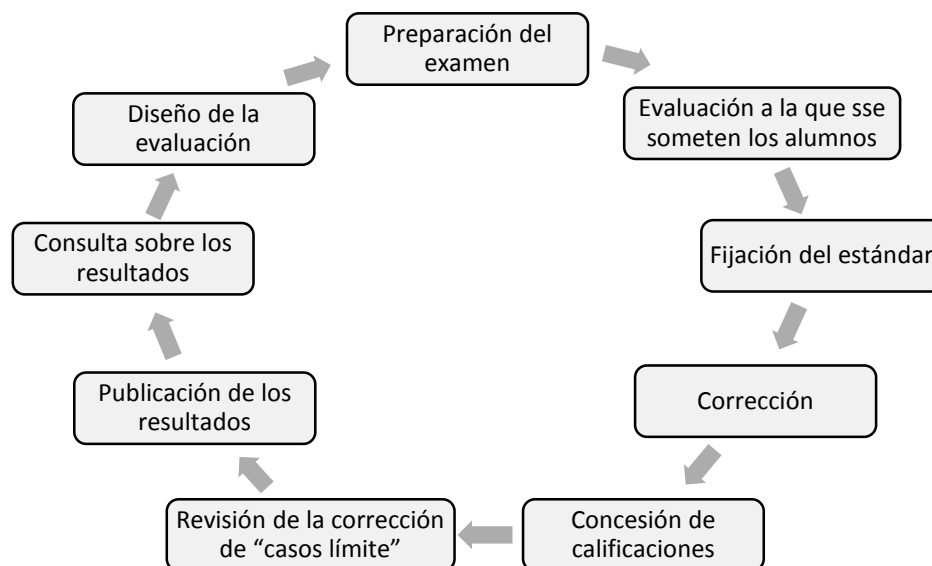


Figura 1. Ciclo de la evaluación.

Cómo se sabe que se trata a todos los alumnos de manera justa en la corrección

A los alumnos pueden evaluarlos sus profesores o un examinador externo. En cualquier caso, el IB debe asegurarse de que se aplican los mismos estándares.

Hay investigaciones que revelan que distintos examinadores pueden tener opiniones levemente distintas sobre la puntuación que merece un trabajo. Para evitar que a los alumnos se les puntúe de manera diferente, el IB nombra como examinador principal a un experto en la asignatura con experiencia. El examinador principal decide qué puntuación se debe conceder, y las demás personas implicadas en la corrección deben seguir ese estándar, de modo que el resultado del alumno no dependa de la persona que corrija su trabajo.

En qué consiste la moderación

La moderación se utiliza en los trabajos de evaluación interna para garantizar que haya un estándar común en todos los colegios. Como resultado de la moderación, las puntuaciones de un colegio pueden subir, bajar o quedarse igual.

El objetivo de la moderación es comprobar con qué precisión y coherencia ha aplicado el profesor los criterios de evaluación al corregir el trabajo de los alumnos.

La muestra de la evaluación interna se selecciona cuidadosamente para garantizar que el rango de puntuación del colegio esté representado adecuadamente. Para la muestra para la moderación, el IB no suele seleccionar a alumnos que hayan logrado la máxima puntuación, ya que así permite que los alumnos con el rango de puntuaciones más elevadas tengan la posibilidad de que sus trabajos se moderen al alza.

El examinador modera la muestra y hace un ajuste a la puntuación del profesor para todos los alumnos del colegio en ese componente; para ello se basa en una comparación estadística entre los dos grupos de puntuaciones. Si el profesor puntúa por exceso o por defecto de manera repetida, se realizará el mismo ajuste en todas las puntuaciones que otorgue el profesor. Si el profesor puntúa por exceso o por defecto en la parte superior o en la parte inferior del rango de puntuaciones, es posible que el ajuste varíe en todo el rango de puntuaciones.

En algunos casos, tal vez no sea posible calcular un ajuste de moderación con el trabajo de muestra entregado, por lo que se solicitarán más trabajos para que se pueda aplicar un ajuste de moderación justo. Por ese motivo, el trabajo de todos los alumnos debe estar disponible hasta que se publiquen los resultados.

El proceso de moderación también se aplica a una pequeña cantidad de componentes de evaluación externa. En dichos casos, cada examinador envía una muestra de su corrección a un examinador supervisor, y las puntuaciones de cada examinador se ajustan para que vayan en consonancia con el estándar de corrección del examinador principal.

Jerarquía de la moderación

La moderación es un proceso jerárquico diseñado para garantizar que las puntuaciones finales que se otorgan a cada colegio vayan en consonancia con el estándar de corrección que establece el examinador principal. A continuación se ilustran los distintos niveles de jerarquía de un componente de evaluación interna con muchos alumnos, en el que los examinadores también están sujetos a la moderación.

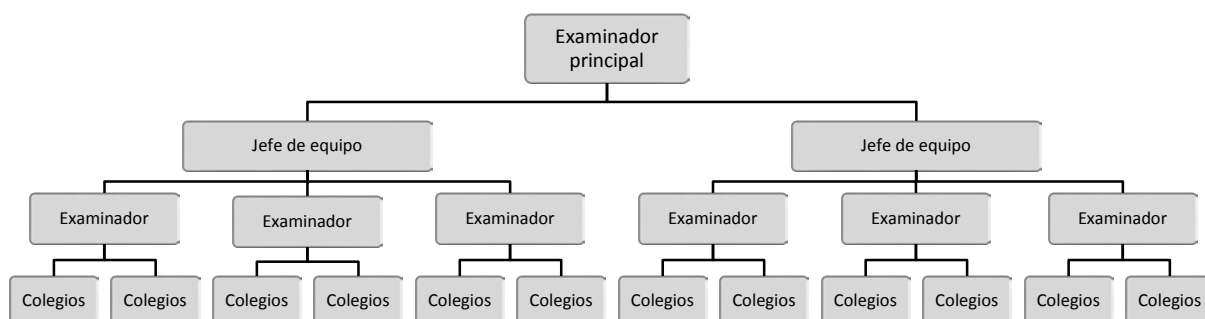


Figura 2. Jerarquía de la moderación.

Este diagrama muestra que:

- Las puntuaciones de un colegio se pueden ajustar en función de la muestra que se entregue a un examinador
- Las puntuaciones de cada examinador se revisan y se ajustan en función de una muestra de su corrección que se entrega a un examinador supervisor (el jefe de equipo)
- A su vez, las puntuaciones del jefe de equipo se pueden ajustar basándose en una muestra de sus correcciones que se entrega al examinador principal

Esta cadena de moderación permite que se hagan una serie de ajustes en las puntuaciones de un colegio antes de otorgar una puntuación moderada final (que siga el estándar del examinador principal). Para los componentes de evaluación externa moderados, la puntuación total que otorga un examinador en una prueba no necesariamente será la puntuación “moderada” final que se le otorgue al alumno. La puntuación final es el resultado que se esperaría que el examinador principal otorgara si él hubiera corregido el trabajo.

Cómo se sabe que los examinadores corrigen el trabajo de los alumnos adecuadamente

Antes de corregir el trabajo, se celebra una reunión de estandarización en la que el examinador principal fija el estándar de corrección de la evaluación; para ello cuenta con la ayuda de examinadores supervisores. Todos los examinadores deben adherirse a ese estándar cuando corrijan. El equipo de estandarización revisa los exámenes con las respuestas de los alumnos y determina un rango de exámenes de práctica, exámenes de aptitud y exámenes de control adecuados para ayudar así a garantizar la fiabilidad de los examinadores durante toda la convocatoria de exámenes.

El esquema de calificación se finaliza como parte del proceso de estandarización a fin de lograr que el proceso de corrección sea más claro, ayudar a que los examinadores sigan el estándar establecido, y contribuir a que haya coherencia en la corrección entre los examinadores.

Corrección electrónica y exámenes de control

Después de que los alumnos se sometan a sus exámenes, sus respuestas se envían a un centro para que las escaneen o se cargan electrónicamente. A continuación, los examinadores tienen acceso a las respuestas a través de una herramienta de corrección en línea que les asigna respuestas aleatoriamente. Las respuestas son anónimas para evitar que los examinadores tengan sesgos.

El proceso que deben seguir los examinadores tiene tres fases: exámenes de práctica, exámenes de aptitud y corrección en condiciones reales. Para controlar la corrección en condiciones reales, se utilizan exámenes de control.

La fase de corrección de exámenes de práctica permite que los examinadores vean exámenes con respuestas de los alumnos que ya ha corregido el examinador principal (es decir, exámenes corregidos definitivos) para saber cómo aplicar el esquema de calificación. Los examinadores reciben comentarios automáticamente para que entiendan con más claridad cómo se debe realizar la corrección.

El proceso de corrección de los exámenes de aptitud exige que los examinadores demuestren que pueden aplicar el esquema de calificación adecuadamente; para ello, se les pone a prueba con un conjunto de exámenes con respuestas de los alumnos que ha seleccionado el examinador principal. Si el examinador no logra demostrar que puede aplicar el esquema de calificación de manera adecuada, no se le permitirá empezar a corregir los trabajos de los alumnos.

Una vez que el examinador reúne los requisitos para empezar con la corrección, se hace un seguimiento de sus correcciones mediante exámenes de control que se introducen aleatoriamente entre los que se le asignan. Un examen de control es un examen que ya ha corregido definitivamente el examinador principal. Los examinadores no saben que están corrigiendo un examen de control, lo que garantiza que estos exámenes de control se corrijan del mismo modo que el resto de exámenes asignados.

A cada asignatura se le asigna un valor de tolerancia para ayudar a determinar que los examinadores hagan sus correcciones según el estándar de corrección que establecen el examinador principal y el equipo de estandarización; por ello, la puntuación del examinador debe encontrarse dentro de este margen de tolerancia. Este margen refleja las diferencias legítimas en las puntuaciones que distintos examinadores otorgan a un mismo trabajo. Imagínese a dos profesores de su colegio que corrigen un trabajo; ambos coinciden en que el trabajo es bueno, pero uno le daría una puntuación de 29 y el otro de 30.

Corrección por preguntas

La corrección por preguntas sigue el mismo proceso de corrección que se ha detallado anteriormente. Sin embargo, facilita la corrección de los exámenes escritos de los alumnos, ya que divide las respuestas en secciones conocidas como *grupo de preguntas* (*question item group*, QIG). Un grupo de preguntas puede estar formado por una pregunta, una sección o un tema. Los examinadores corrigen estas secciones más reducidas del mismo modo que los exámenes completos (es decir, primero corrigen las partes de práctica y de aptitud antes de poder pasar a los exámenes en condiciones reales con exámenes de control). A los examinadores se les hace un seguimiento por separado en cada grupo de preguntas, y cada uno de estos grupos tiene su propio margen de tolerancia en la corrección.

Al dividir los exámenes en grupos de preguntas, los examinadores pueden corregir los temas en los que son especialistas y centrarse en corregir pregunta a pregunta, lo cual aumenta la fiabilidad de la corrección.

Ajuste proporcional

Para combinar las puntuaciones y los límites de calificación de los distintos componentes y obtener una puntuación total en forma de porcentaje, lo primero que hay que hacer es ajustar proporcionalmente las puntuaciones. Estas puntuaciones ajustadas significan que cada componente tiene un valor distinto, que depende de cómo contribuya al total de la evaluación en la asignatura.

Por ejemplo: en el Nivel Medio de la asignatura de Geografía del Programa del Diploma, la puntuación máxima de la evaluación interna es de 30, pero representa un 25 % del total de la evaluación. Por tanto, a la puntuación de la evaluación interna se le aplica un ajuste proporcional de 25/30, es decir, se reduce para que el máximo sea de 25 puntos y no de 30. La prueba 1 tiene una puntuación máxima de 60, pero se le asigna un 40 % del total de la evaluación, lo que significa que la puntuación se reduce de 60 a 40 (ajuste

proporcional de 40/60). La prueba 2 tiene una puntuación máxima de 40, pero se le asigna un 35 % del total de la evaluación, lo que significa que la puntuación se reduce de 40 a 35 (ajuste proporcional de 35/40). Por lo tanto, la puntuación total disponible para la asignatura es de $30 + 60 + 40 = 130$, pero después de aplicar el ajuste proporcional, la puntuación máxima ajustada será de 100.

Componente	Puntuación máxima	Porcentaje del total de la evaluación	Factor de ajuste proporcional
Evaluación interna	30	25 %	25/30
Prueba 1	60	40 %	40/60
Prueba 2	40	35 %	35/40

A continuación, esa puntuación ajustada se utiliza para determinar la calificación final que se otorga al alumno, y para ello se utilizan los límites de calificación de la asignatura.

Cómo se determinan los límites de calificación

Hay una diferencia entre *corregir* y *calificar* el trabajo de un alumno.

Al corregir, se reconoce la labor que ha realizado un alumno y se le otorgan unas puntuaciones según un esquema de calificación o un marco similar. Esto representa en qué medida el alumno realizó correctamente la tarea de evaluación. La puntuación en sí misma no significa nada más.

A la hora de decidir la calificación, se juzga la calidad del trabajo del alumno utilizando como referencia un estándar definido. Para ello, se tienen en cuenta la dificultad de la tarea y también la proporción de ella que se completó. Así pues, la calificación tiene significado y pertinencia y está pensada para poder comparar el desempeño en otras evaluaciones.

Se debe recordar que: Las *puntuaciones* representan cuánto ha avanzado un alumno, mientras que las *calificaciones* tienen en cuenta lo complicado que era el trayecto.

La concesión de calificaciones es el proceso mediante el cual el IB convierte las puntuaciones en calificaciones, y sucede poco antes de la publicación de los resultados. Los examinadores supervisores (incluido el examinador principal) y determinados miembros del personal del IB son los encargados de establecer los límites de calificación; además, también se encargan de resolver las cuestiones pendientes de los exámenes. Los límites de calificación se determinan mediante una combinación de valoraciones, pruebas estadísticas, descriptores de calificaciones finales y los formularios G2 que entregan los profesores. Los formularios G2 proporcionan información importante para los examinadores supervisores, y les ayudan a acordar los límites de calificación teniendo en cuenta las percepciones que los profesores tienen acerca de los exámenes.

Cómo se sabe que los resultados son correctos

Debido al uso de exámenes de aptitud y exámenes de control, el IB tiene la confianza de que todos los exámenes se corrigen adecuadamente. Sin embargo, es importante que los alumnos reciban la calificación correcta en la asignatura en función del trabajo que entreguen, por lo que hay controles de calidad adicionales sobre la corrección del trabajo de los alumnos.

Revisión de la corrección de “casos límite”

La intención del proceso de los “casos límite” es identificar a los alumnos cuyas calificaciones se arriesgan a ser demasiado bajas teniendo en cuenta las calificaciones previstas, y revisar la corrección de estos. También sirve para detectar los problemas de corrección que no se hayan identificado anteriormente. Esta práctica garantiza que los alumnos que se encuentran en el límite reciban una calificación precisa.

Consulta sobre los resultados

Si por cualquier motivo algún alumno o el colegio creen que la calificación otorgada es incorrecta, se puede solicitar una consulta sobre los resultados. Tras la publicación de los resultados, el coordinador puede solicitar una consulta sobre los resultados de:

- Categoría 1: revisión de la corrección de todos los componentes de evaluación externa de un alumno en una asignatura.
- Categoría 2: copias de material de algún componente de evaluación externa.
- Categoría 3: revisión de la moderación de la muestra de un componente de evaluación interna de un colegio.
- Informe de categoría 1: informe sobre la revisión de la corrección de la consulta sobre los resultados de categoría 1.

Se debe abonar una tasa por cada una de las categorías anteriores, salvo cuando se modifica una calificación final como consecuencia de una revisión de la corrección de categoría 1.

Un examen devuelto puede contener comentarios útiles del examinador al que se le había asignado. Sin embargo, no se puede garantizar que sea así, ya que los examinadores no están obligados a escribir comentarios cuando corrigen el trabajo de los alumnos. Consulte la siguiente sección si desea obtener más información sobre cómo recibir comentarios.

Tal como se comenta en la sección “En qué consiste la moderación”, la puntuación que otorga un examinador en una prueba no necesariamente será la puntuación “moderada” final; la puntuación que se da al alumno en el sistema de información del IB (IBIS) es la correcta.

Comentarios disponibles

Aproximadamente una semana después de publicar los resultados, se publican en IBIS todos los informes sobre la evaluación interna. El informe sobre la evaluación interna representa una oportunidad para que los examinadores realicen comentarios sobre la idoneidad de las tareas y sobre cualquier cuestión administrativa. Se aconseja a los examinadores sujetos a la moderación que no hagan comentarios específicos sobre las puntuaciones que haya otorgado el profesor de la muestra de la evaluación interna, ya que no saben si también se aplicará un factor de moderación adicional (al propio examinador). Los examinadores que no están sujetos a la moderación pueden hacer comentarios más detallados sobre los ajustes que han realizado a la puntuación del profesor.

Los examinadores que corrigen componentes de evaluación externa no están obligados a escribir comentarios al corregir el trabajo de los alumnos, ya que lo importante es que los examinadores corrijan de manera adecuada, más que aportar notas y recomendaciones sobre el trabajo. Los comentarios sobre el desempeño del alumno conforman una parte importante de la evaluación formativa que realiza el profesor a lo largo del curso.

El Centro pedagógico en línea (CPEL) es otra fuente de comentarios donde se pueden encontrar tanto informes generales de las asignaturas como material de ayuda al profesor.

Preparación de los exámenes

Para poder afrontar el desafío de preparar exámenes válidos, el trabajo de preparación de un conjunto de exámenes puede empezar hasta dos años antes de que los alumnos se sometan a ellos. Como parte de este proceso, se mantiene una reunión de edición de exámenes en la que participan el responsable de la asignatura, varios examinadores principales del componente y una selección de otros examinadores. Todos ellos discuten sobre los exámenes y los esquemas de calificación antes de finalizar la evaluación, lo cual garantiza que el estándar de evaluación se encuentra en el nivel correcto. Los distintos exámenes para cada nivel se revisan en conjunto para determinar en qué medida se cubren el contenido del curso y sus objetivos. Además, el grupo revisa de manera crítica cada pregunta, fijándose en su idoneidad, su grado de precisión y las posibles ambigüedades o sesgos. Este grupo se ocupará de que los esquemas de calificación sean claros, concisos y completos, para que los puedan seguir los examinadores.

Qué se hace para mejorar la calidad de la evaluación

El proceso de evaluación es un ciclo continuo, por lo que la experiencia cosechada en convocatorias anteriores contribuye al desarrollo del diseño de la evaluación y a la preparación de exámenes de futuras convocatorias. El IB continuará progresando en el desarrollo de mecanismos de control de calidad para que las evaluaciones ofrezcan resultados válidos y fiables para los alumnos. Para obtener información más detallada acerca de la evaluación en el IB, se puede visitar el CPEL y el sitio web del IB.