

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 realizan los alumnos sean significativas y justas.

Trata de los siguientes temas:

- Cómo mantener la coherencia en las correcciones
- Cómo mejorar la calidad de la corrección
- Principios relativos a las consultas 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 correctos.

El ciclo de evaluación es un proceso continuo en el que cada etapa recibe la influencia de la anterior e influye en la siguiente.



Figura 1. Ciclo de la evaluación

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

A los alumnos pueden evaluarlos sus profesores o un examinador externo. El IB debe asegurarse de que se aplican los mismos estándares en ambos casos.

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 cómo se deben conceder las puntuaciones, y las demás personas que participen 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 se aplique 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 para la evaluación interna se selecciona cuidadosamente para garantizar que el rango de puntuaciones del colegio esté representado adecuadamente. Para la muestra que se somete a moderación, el IB no suele seleccionar a alumnos que hayan logrado la máxima puntuación, ya que así permite que los alumnos que estén dentro del rango de puntuaciones más elevado tengan la posibilidad de que sus puntuaciones suban a raíz de la moderación.

Actualmente, el IB utiliza un proceso de moderación llamado **toma dinámica de muestras**.



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 concede puntuaciones excesivas o demasiado bajas 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 a partir de los trabajos que forman parte de la muestra entregada, 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 cierre la convocatoria.

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

Antes que se inicie la corrección, 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 ceñirse a ese estándar cuando corrijan. El equipo de estandarización revisa exámenes con las respuestas de los alumnos y determina un rango de respuestas de práctica, de aptitud y de control adecuado que ayude a garantizar la fiabilidad de la corrección que realicen los examinadores durante toda la convocatoria de exámenes.

El esquema de calificación se concreta durante el 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 todos ellos.

Corrección electrónica y sistema de control

Después de que los alumnos realizan los 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: corrección de respuestas de práctica, corrección de respuestas de aptitud y corrección en condiciones reales. Para controlar la corrección en condiciones reales, se utilizan respuestas de control.

La fase de corrección de respuestas de práctica permite que los examinadores vean respuestas que ya ha corregido el examinador principal (es decir, respuestas corregidas definitivas) para saber cómo aplicar el esquema de calificación. Los examinadores reciben comentarios automáticamente para aclararles cómo se debe realizar la corrección.

El proceso de corrección de respuestas 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 respuestas 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 respuestas de control que se introducen aleatoriamente entre las que se le asignan. Una respuesta de control es una respuesta que ya ha corregido definitivamente el examinador principal. Los examinadores no saben que están corrigiendo una respuesta de control, lo que garantiza que corrijan este tipo de respuestas del mismo modo que las demás que se les han asignado.

A cada asignatura se le asigna un valor de tolerancia para ayudar a determinar que los examinadores están corrigiendo según el estándar de corrección que establecen el examinador principal y el equipo de estandarización. La puntuación del examinador debe encontrarse dentro de este margen de tolerancia. El margen de tolerancia refleja las diferencias legítimas en las puntuaciones que distintos examinadores otorgan a un mismo trabajo. Imaginemos, por ejemplo, a dos profesores de un colegio que corrigen el mismo 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 que se ha detallado anteriormente. Sin embargo, facilita la corrección de los exámenes, ya que divide las respuestas en secciones conocidas como **grupos de preguntas** (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 las pruebas completas (es decir, corrigen las respuestas de práctica y de aptitud antes de corregir respuestas en condiciones reales, que incluyen respuestas 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.

Al dividir los exámenes en grupos de preguntas, los examinadores pueden corregir los temas en los que son especialistas y centrarse en corregir pregunta por 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 proporcionalmente significan que cada componente representa un porcentaje distinto, que depende de cómo contribuya al total de la evaluación en la asignatura.

Por ejemplo: en el Nivel Medio de la asignatura de Biología del Programa del Diploma, la puntuación máxima de la evaluación interna es de 24, pero representa un 20 % del total de la evaluación. Por tanto, a la puntuación de la evaluación interna se le aplica un ajuste proporcional, multiplicándola por un factor de 20/24, para expresarla en una escala en la que el máximo sea de 20 y no de 24. El componente de la

prueba 1 tiene una puntuación máxima de 30 y constituye un 20 % del total de la evaluación, lo que significa que la puntuación se ajusta proporcionalmente multiplicándola por un factor de 20/30. El componente de la prueba 2 tiene una puntuación máxima de 50 y constituye un 40 % del total de la evaluación, lo que significa que la puntuación se ajusta proporcionalmente multiplicándola por un factor de 40/50. El componente de la prueba 3 tiene una puntuación máxima de 35 y constituye un 20 % del total de la evaluación, lo que significa que la puntuación se ajusta proporcionalmente multiplicándola por un factor de 20/35). Por lo tanto, la puntuación total disponible para la asignatura es de $24 + 30 + 50 + 35 = 139$, 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	24	20 %	20/24
Prueba 1	30	20 %	20/30
Prueba 2	50	40 %	40/50
Prueba 3	35	20 %	20/35

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, se encargan de resolver las cuestiones pendientes de los exámenes. Los límites de calificación se determinan utilizando 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 límites de calificación adecuados 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 respuestas de aptitud y 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 en riesgo

La intención del proceso de revisión de la corrección de los casos en riesgo es identificar a los alumnos cuyas calificaciones están en riesgo de ser demasiado bajas teniendo en cuenta las calificaciones previstas, y revisar la corrección de estos. Este proceso se realiza antes de publicar los resultados. También sirve para detectar los problemas de corrección que tal vez no se hayan identificado anteriormente. Esta práctica garantiza que los alumnos que se encuentran en el límite reciban una calificación correcta.

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 realizar estos tipos de consultas sobre los resultados:

- 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 realizada a partir de una consulta sobre los resultados de categoría 1.

Se debe abonar una tasa por cada una de las categorías anteriores, salvo cuando se modifique una calificación final como consecuencia de una revisión de la corrección realizada a partir de una consulta sobre los resultados 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.

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.

Los examinadores que corrigen componentes de evaluación externa no están obligados a escribir comentarios al corregir el trabajo de los alumnos, pues en lo que se hace hincapié es en que los examinadores se aseguren de corregir de manera adecuada, y no en que aporten comentarios 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 de recursos para los programas es otra fuente de comentarios donde se pueden encontrar tanto informes generales de las asignaturas como material de ayuda al profesor.

Cómo se crean los exámenes

Para poder afrontar el desafío de crear exámenes válidos, el trabajo de preparación de un conjunto de exámenes puede empezar hasta dos años antes de que los alumnos los realicen. Como parte de este proceso, se mantiene una reunión de edición de pruebas de evaluación 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 que produzca la versión final de las evaluaciones, 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 ocupa 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 influye en el diseño de la evaluación y 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 pueden visitar el Centro de recursos para los programas y el sitio web del IB.