

# Principios y prácticas de evaluación del IB: guía sobre la evaluación para examinadores

## Introducción

Este documento es una guía resumida del ciclo de evaluación y está destinado a los examinadores del IB. Los examinadores desempeñan una función fundamental en el proceso de evaluación, y esta guía está pensada para explicar cómo sus contribuciones afectan a todo el ciclo de evaluación. Para obtener información más detallada acerca de la evaluación en el IB, visite el sitio web del IB.

Las evaluaciones sumativas deben proporcionar a los alumnos resultados *significativos*, *justos* y *fiabes*:

- **Significativos** porque las evaluaciones deben medir aquello que se pretendía ponderar, tal como se indicara en la guía de la asignatura.
- **Justos** porque las evaluaciones no deben favorecer a ningún grupo concreto de alumnos.
- **Fiabes** porque los alumnos deben recibir los mismos resultados independientemente de quién corrija sus evaluaciones o en qué día lo hagan.

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.

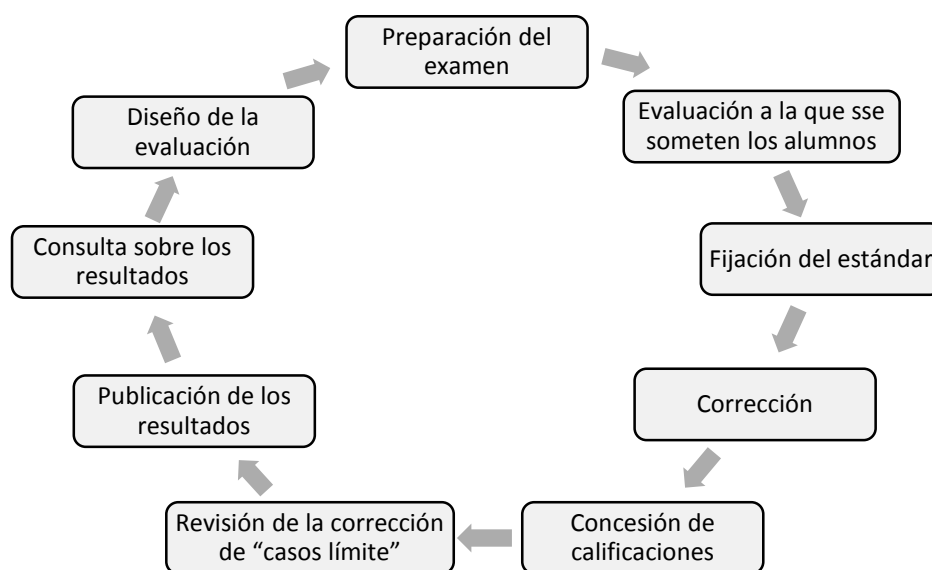


Figura 1. Ciclo de la evaluación.

## Creación y preparación de exámenes: ¿quién redacta las evaluaciones?

El proceso de creación de los exámenes empieza aproximadamente entre 18 meses y 2 años antes de que los alumnos se sometan a los exámenes. Los miembros del equipo de examinadores supervisores son los responsables de preparar uno o más exámenes para componer el conjunto de exámenes necesarios para una asignatura. Los comentarios que usted proporcione sobre los exámenes y los esquemas de calificación resultarán muy importantes para mejorar constantemente. Por ejemplo, si en la convocatoria de exámenes anterior el equipo supervisor recibió comentarios en los que se indicaba que determinados tipos de preguntas no eran accesibles para los alumnos o resultaban complicadas de corregir de manera fiable, esta información ayudará a diseñar mejores exámenes en futuras convocatorias.

## ¿Saca el máximo partido a la capacitación para examinadores?

Los examinadores pueden recibir capacitación para la amplia mayoría de las asignaturas. Esto les permitirá familiarizarse con la evaluación que van a corregir y les ayudará a entender cómo deben realizar la corrección. Al recibir la capacitación, estarán mejor preparados para las fases de corrección de exámenes de práctica y de aptitud de la convocatoria de exámenes.

## ¿Entiende el estándar de corrección del examinador principal?

El examinador principal decide cuál es la “respuesta correcta” y también cómo se conceden las puntuaciones. Todos los examinadores deben seguir las indicaciones del examinador principal cuando corrijan.

Los resultados de los alumnos son muy importantes y es esencial que el alumno reciba el mismo resultado que habría obtenido si un examinador principal hubiera realizado la corrección, independientemente de quién sea el examinador que corrija su trabajo. Por ese motivo, el IB dedica grandes esfuerzos para garantizar que la corrección del trabajo del alumno sea fiable y precisa. Actualmente existen dos métodos distintos para garantizar que todos los examinadores utilicen el mismo estándar a la hora de corregir un componente: el sistema de exámenes de control y la moderación. La mayoría de componentes de evaluación externa ahora se corrige en pantalla, y se utiliza el método de exámenes de control para asegurarse de que todos los examinadores sigan el mismo estándar. El otro método utilizado es la moderación, que se explica más adelante en esta guía.

En el método de exámenes de control, todos los examinadores deben aprender el estándar de corrección del examinador principal. Esto se consigue mediante un esquema de calificación o notas para la corrección, capacitación y práctica.

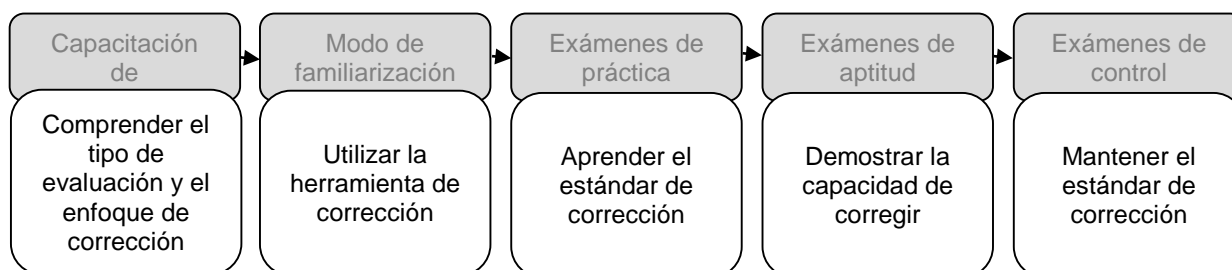


Figura 2. Esquema del sistema de exámenes de control.

Después de realizar la evaluación y de recibir varias respuestas de los alumnos, el examinador principal mantiene una reunión de estandarización con los demás examinadores supervisores. Los examinadores supervisores deben acordar las puntuaciones para varios exámenes que se usarán para definir el modelo de calidad. Dichos exámenes se denominan *exámenes corregidos definitivos*, y pueden ser exámenes de práctica, de aptitud y de control. Al corregir un examen corregido definitivo, las puntuaciones que usted otorgue se comparan con las puntuaciones acordadas, con el fin de medir la precisión de la corrección que usted ha realizado. Esto sirve para saber en qué medida los alumnos recibirán resultados fiables en los trabajos que usted les corrija.

Otro propósito de la reunión es observar trabajos reales de alumnos y ofrecer aclaraciones a los examinadores, como ejemplos adicionales de respuestas aceptables en las notas para la corrección.

Antes de empezar a corregir los exámenes de práctica, debe leer con atención los esquemas de calificación o las notas adicionales que se le envíen, ya que esta fase del proceso le da la oportunidad de entender mejor el estándar de corrección que se exige en esa convocatoria. Cuando corrija los exámenes de práctica, puede comparar las puntuaciones que otorgue con las puntuaciones definitivas y leer los comentarios o las anotaciones que el examinador principal haya escrito en los exámenes. También puede contactar con el jefe de equipo para hacerle las preguntas que tenga.

## Cómo le afecta el modelo de calidad

Es importante que, al acabar de corregir los exámenes de práctica, esté seguro de que entiende el esquema de calificación, los criterios de evaluación y las notas para la corrección, y que comprenda cómo ponerlos en práctica en los exámenes de la convocatoria en curso. Debe entender con claridad cómo y porqué se otorgaron las puntuaciones definitivas de los exámenes de práctica. Llegado a este punto, puede empezar con el proceso de corrección de exámenes de aptitud.

La primera fase de corrección de los exámenes de aptitud representa una oportunidad para demostrar que puede corregir siguiendo el estándar fijado; por ese motivo, su jefe de equipo no le puede ayudar. Tanto las puntuaciones totales como las que usted otorgue a cada pregunta se comparan con las puntuaciones definitivas. Para poder pasar a la corrección en condiciones reales, debe corregir todos los trabajos dentro de un margen de tolerancia predeterminado. Se aplica un margen de tolerancia porque se acepta que puede haber diferencias legítimas en las puntuaciones que distintos examinadores otorguen a un mismo trabajo; sin embargo, es muy importante reducir al mínimo estas diferencias para que afecten de la menor manera posible a los resultados de los alumnos. Si las puntuaciones que ha otorgado se encuentran fuera del margen de tolerancia en un examen o más, recibirá comentarios del jefe de equipo para que entienda cómo se deberían haber otorgado las puntuaciones.

Cuando el jefe de equipo esté seguro de que usted ha comprendido cómo conceder las puntuaciones, le dará acceso a un segundo grupo de exámenes de aptitud. Si logra corregir todos estos exámenes dentro del margen de tolerancia, se le permitirá que empiece la corrección en condiciones reales. Superar la fase de corrección de los exámenes de aptitud sirve para demostrar que puede aplicar con precisión el esquema de calificación y que es capaz de corregir en consonancia con el examinador principal, por lo que los alumnos recibirán resultados justos. Es entonces cuando se considera que está preparado para pasar a la corrección en condiciones reales. A pesar de todo, revise las diferencias que hay entre las puntuaciones que usted otorgue y las puntuaciones definitivas; de ese modo, perfeccionará su capacidad de corrección y mejorará la fiabilidad de los resultados de los alumnos.

Cuando corrija en condiciones reales, recibirá exámenes de control en intervalos regulares para comprobar que continúa aplicando adecuadamente el esquema de calificación y que se rige por el estándar de corrección que se ha establecido. Si corrige un examen de control dentro del margen de tolerancia, podrá ver las puntuaciones definitivas y los comentarios para identificar en qué áreas su puntuación difiere de la definitiva, lo que le ayudará a perfeccionar su corrección a fin de que se asemeje lo más posible a la del examinador principal; de esta manera, se garantizan resultados fiables para los alumnos. Si corrige un examen de control fuera del margen de tolerancia, antes de continuar con la corrección deberá revisar las puntuaciones definitivas y las anotaciones para entender por qué su corrección no ha sido adecuada. Si continúa corrigiendo los exámenes de control fuera del margen de tolerancia, no podrá seguir corrigiendo. De este modo, se evita que los alumnos reciban puntuaciones inadecuadas y que no sean fiables.

## Corrección por grupos de preguntas

En algunos componentes, el examen se divide en secciones más cortas que se corrigen independientemente. Cada una de estas secciones recibe el nombre de *grupo de preguntas* (*question item group*, QIG). Al dividir el examen de esta manera, puede utilizar una parte más pequeña del esquema de calificación, lo que reduce a su vez la carga cognitiva. Es posible que algunos grupos de preguntas le resulten más sencillos o más amenos de corregir que otros. Los exámenes también se pueden dividir por opción o por tema, lo que permite que corrija las opciones o los temas con los que se sienta más cómodo.

La corrección por grupos de preguntas se hace en pantalla, por lo que el modelo de calidad que se usa es el de la corrección de exámenes de control. Tendrá que finalizar la corrección de los exámenes de práctica y de aptitud de cada grupo de preguntas que quiera corregir. Durante la corrección en condiciones reales, se hace un seguimiento de su progreso mediante exámenes de control.

## Moderación de la evaluación interna

En ocasiones, resulta más adecuado que sea un profesor y no un examinador quien juzgue algunas habilidades, como las que se evalúan con el paso del tiempo. Los examinadores moderan externamente los componentes de evaluación interna, y para ello utilizan muestras de los trabajos de los alumnos. Los examinadores que moderan evaluaciones internas normalmente se denominan *moderadores*, para dejar claro que moderan trabajos ya corregidos en lugar de examinar trabajos que aún no se han sometido a

corrección. El sistema de información del IB (IBIS) selecciona las muestras, que deben ser representativas del rango de puntuaciones que haya en cada colegio.

Todos los componentes de evaluación interna se corrigen aplicando criterios de evaluación o de bandas de puntuación. La mayoría de veces el profesor tiene acceso a más cantidad de información que el examinador sobre el contexto y el proceso que rodea al trabajo de un alumno. Por ese motivo, al moderar se pide que se juzgue si la corrección del profesor resulta adecuada; no se trata de volver a corregir el trabajo y rechazar las puntuaciones del profesor. De hecho, esas puntuaciones solamente se pueden modificar si está seguro de que resultan inadecuadas.

Las puntuaciones que usted otorga y las del profesor se comparan mediante una regresión lineal y se aplica un ajuste adecuado en cada colegio para que las puntuaciones del profesor estén en consonancia con las que usted haya concedido. Este ajuste se denomina *factor de moderación*. Es posible que además tenga que enviar una muestra a su jefe de equipo. En ese caso, se compararán las puntuaciones de la misma manera, para comprobar si es necesario hacer algún ajuste más y que sus puntuaciones vayan en consonancia con el estándar del jefe de equipo. A su vez, el jefe de equipo enviará al examinador principal una muestra de sus propias correcciones. Este proceso de moderación se hace sin que usted sea consciente de los ajustes que se hacen en sus puntuaciones cuando escriba el informe sobre la evaluación interna de cada colegio, así que no debe realizar ningún comentario sobre las diferencias entre sus puntuaciones y las del profesor.

## Toma dinámica de muestras para moderar en pantalla

Para los trabajos de evaluación interna que se corrigen electrónicamente se utiliza otro modelo de calidad que combina características de la moderación y del uso de exámenes de control. Al principio, moderará los trabajos de tres alumnos del total de muestras de cada colegio. Si las diferencias entre las puntuaciones del profesor y las suyas se encuentran dentro de un margen de tolerancia predeterminado, se considerará que la corrección del profesor es adecuada y que no es necesario ajustar las puntuaciones. Si alguna de las puntuaciones está fuera del margen de tolerancia, se le dará el resto de muestras para que las modere, y se calculará un factor de moderación adecuado que se aplicará a las puntuaciones del profesor.

En el modelo de toma dinámica de muestras, deberá superar un proceso de aptitud similar al empleado para los componentes de evaluación externa que se corrigen electrónicamente. Se hará un seguimiento mediante exámenes de control en vez de enviar una muestra de sus correcciones a un examinador supervisor. Puesto que usted corrige empleando el mismo estándar que el examinador principal, no se aplicará ningún ajuste de moderación a las puntuaciones que otorgue. Esto permite que pueda escribir comentarios detallados y útiles en el informe sobre la evaluación interna, incluidos comentarios sobre las diferencias entre sus puntuaciones y las del profesor.

## Moderación de la evaluación externa

A los pocos componentes de evaluación externa que todavía no se corrigen de manera electrónica se les aplica el modelo de calidad de la moderación. El proceso es similar al de la moderación de la evaluación interna que se ha descrito anteriormente.

El proceso de moderación confía en que sus correcciones sean coherentes a lo largo del período de corrección. Por este motivo, a diferencia de lo que sucede en el método de exámenes de control, usted no recibirá comentarios sobre las puntuaciones que otorgue hasta que acabe todas las correcciones (incluida la revisión de la corrección que realice si hay consultas sobre los resultados). Si su muestra revela que no ha sido coherente en la corrección o que no ha aplicado adecuadamente los criterios de evaluación o las bandas de puntuación, no se podrá aplicar un ajuste de moderación y otro examinador volverá a corregir los trabajos que tenía usted asignados para garantizar que los alumnos reciban resultados justos.

# Cómo se establecen los límites de calificación

En el IB, *corregir* y *calificar* los trabajos de los alumnos son dos acciones distintas.

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 indica 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 puntuación otorgada. Las calificaciones tienen un significado y una pertinencia concretos, y están pensadas para que se puedan comparar con las calificaciones en las demás asignaturas.

**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.

Al final del período de corrección, se celebra una reunión de evaluación para cada asignatura con el objetivo de que los examinadores supervisores revisen cada componente y establezcan límites de calificación junto con el responsable de la asignatura. Antes de la reunión, los examinadores supervisores y los responsables de la asignatura pueden ver los comentarios de los profesores sobre cada componente que entregaron en los formularios G2; estos comentarios ofrecen información de referencia importante para la reunión. Al completar el informe del examinador, usted también proporciona a la reunión información de referencia importante en la que se comenta el desempeño de los alumnos que le han asignado. Los informes del examinador no solo sirven a la hora de conceder calificaciones, sino que también contribuyen al informe general de la asignatura que se pone a disposición de los colegios.

A principio de curso se fijan los límites de calificación para los componentes de evaluación interna y para los componentes que no son exámenes (trabajo de clase); como la tarea no cambia, los límites de calificación no varían necesariamente tras cada convocatoria, sino que se revisan para comprobar que sigan siendo adecuados. Sin embargo, en cada convocatoria se establecen nuevos límites para cada componente de examen, porque se prepara un examen nuevo para cada convocatoria. Para determinar dónde establecer los límites, se utilizan el juicio profesional de los examinadores supervisores y la información estadística disponible. Si se detecta un cambio significativo del desempeño de los alumnos en un componente o una asignatura, y ese cambio se debe, por ejemplo, a que el examen de la convocatoria vigente sea más difícil que el de la convocatoria anterior, los límites de calificación cambiarán en consecuencia. El equipo de concesión de calificaciones debe garantizar que se siga utilizando el mismo estándar de calificación que en las convocatorias anteriores.

## Revisión de la corrección de “casos límite”

Después de fijar los límites de calificación, se realizan algunos controles adicionales sobre la corrección para garantizar que las calificaciones que se conceden a los alumnos sean justas y correctas. Uno de estos controles es la revisión de la corrección de “casos límite”. Se trata de alumnos con los que se corre el riesgo de que la calificación que han recibido de la asignatura sea baja por motivos incorrectos. En algunos casos, la revisión de la corrección de “casos límite” se centra en los alumnos cuyas calificaciones están dos puntos o más por debajo de lo que estaba previsto, o que necesitan un punto o dos más para obtener una calificación más alta. Otras veces, la revisión de la corrección de “casos límite” se centra en el trabajo de los examinadores cuyas correcciones se han identificado como potencialmente problemáticas. Esta revisión de la corrección de “casos límite” no se hace para identificar si se pueden conseguir algunos puntos más, sino para comprobar que la puntuación otorgada es adecuada; de este modo, se garantiza que los alumnos reciban calificaciones correctas.

## ¿Se es más generoso en las consultas sobre los resultados?

Si los colegios consideran que las puntuaciones que se han otorgado a sus alumnos son incorrectas, pueden solicitar consultas sobre los resultados.

Una consulta sobre los resultados de categoría 1 es una revisión de la corrección de los componentes de evaluación externa de un alumno en una asignatura. La revisión de la corrección suele hacerla un

examinador supervisor. De modo parecido a lo que sucede con la revisión de la corrección de “casos límite”, al revisar la corrección del trabajo de un alumno a raíz de una consulta sobre los resultados, el examinador no pretende conceder puntos adicionales al alumno, sino comprobar la corrección anterior para enmendar los errores que encuentre. Si no se encuentran errores, se considera que la puntuación otorgada es correcta. Si el examinador identifica errores en la corrección que se había realizado, deberá enmendar esa puntuación e introducir un comentario para explicar el cambio en las puntuaciones o proporcionar detalles sobre el error detectado. El IB consultará dicho comentario cuando investigue la puntuación original que le pareció equivocada a la persona encargada de revisarla.

Los colegios también pueden solicitar que se devuelva el trabajo del alumno a través de una consulta sobre los resultados de categoría 2. Si bien no es necesario escribir en los exámenes comentarios para los profesores, es importante que los comentarios que decida hacer sean adecuados y se refieran a los criterios de evaluación, las bandas de puntuación y los esquemas de calificación.

Una consulta sobre los resultados de categoría 3 es una revisión de la moderación de la muestra de evaluación interna de un colegio. Este tipo de consulta sobre los resultados solo está disponible para los colegios en los que haya habido una reducción significativa en las puntuaciones de un profesor. La revisión de la moderación suele hacerla un examinador supervisor. Cuando se revisa la moderación de un colegio es importante que el examinador utilice el mismo estándar que se utilizó en esa convocatoria. Si un examinador realiza cambios en las puntuaciones del examinador original, resulta útil escribir un comentario en la pantalla de la muestra. El IB consultará dicho comentario cuando investigue la moderación original que le pareció inadecuada a la persona encargada de revisarla.