

IB assessment principles and practices: A guide to assessment for examiners

Introduction and overview

This document is intended as a summary guide to the assessment cycle for IB examiners. Examiners are critical to the process of assessment, and this guide is intended to explain how your contributions affect the whole assessment cycle. More detailed information about IB assessment can be obtained from the IB website.

Summative assessments need to provide *meaningful*, *fair* and *reliable* outcomes for candidates:

- **meaningful** because assessments should measure what was intended to be measured, as set out in the subject guide
- **fair** because the assessments should not favour particular groups of candidates
- **reliable** because a candidate should receive the same outcome regardless of who marked their assessment or on what day.

The assessment cycle is a continuous process, whereby each stage is informed by the previous stage and leads into the next stage.

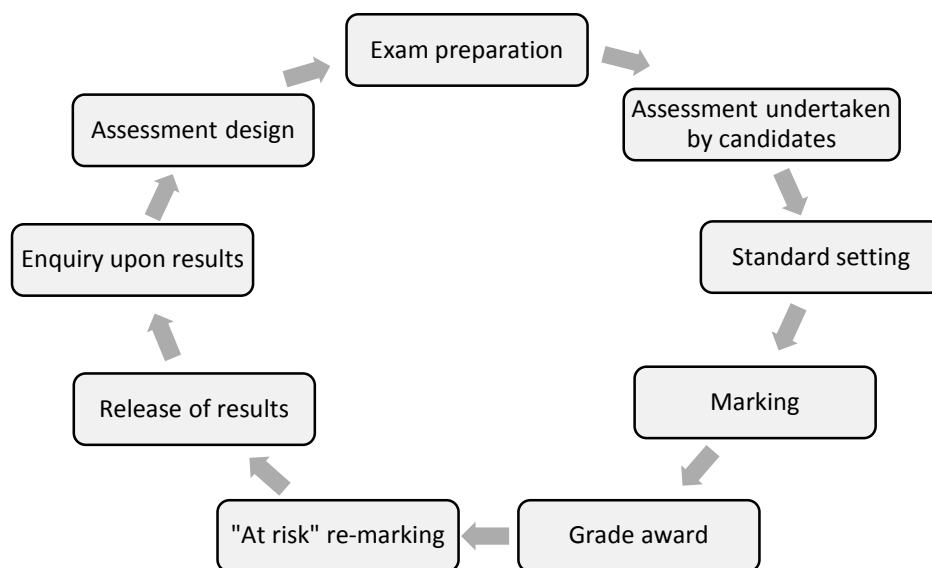


Figure 1. Assessment cycle

Exam authoring and preparation – who writes the assessments?

The process of creating the examinations begins approximately 18 months to two years before they are taken by candidates. Members of the senior examining team are given the responsibility of preparing one or more examinations to make up the complete set that is required for a subject. Your feedback on the examinations and markschemes is critical to continuous improvement. For example, if in the previous examination session the senior team received feedback that certain types of questions were not accessible to candidates or were difficult to mark reliably, this information will be used to help design better examinations for future sessions.

Examiner training – are you making the most of it?

Training is offered to examiners in the vast majority of subjects. It will help familiarize you with the assessment you are going to be marking and help you to understand how you are expected to mark. By completing the training, you will be better prepared for the practice and qualification stages of the examination session.

Marking – do you understand the PE’s standard?

The PE decides what the “right answer” is and how to give marks. Every other examiner must follow the views of the PE when marking.

The outcomes for candidates are of primary importance and it is vital that, whichever examiner marks their work, candidates should receive the same outcome as if it had been marked by the principal examiner. For this reason the IB goes to great lengths to ensure candidate work is marked reliably and accurately. Currently two different methods are used to ensure all examiners mark at the same standard within a component: seeding and moderation. The majority of externally assessed components are now marked on screen, and the seeding method is used to ensure the same standard is applied by all examiners. The other method used is moderation, which is outlined later in this guide.

In the seeding model, all examiners are required to learn the principal examiner’s standard of marking. This is achieved through the markscheme/marketing notes, training and practice.

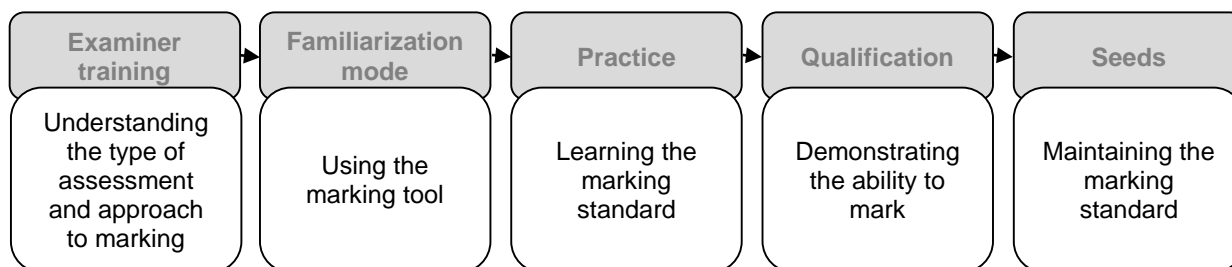


Figure 2. Seeding model

After the assessment has been taken and a number of candidate responses are received, the principal examiner holds a standardization meeting with the other senior examiners. The senior examiners will agree on the marks for a number of scripts that will be used to set up the quality model. These are called definitively marked scripts and are divided into practice, qualification and seeds. When you mark a definitively marked script, your marks are compared to the agreed marks to measure how accurate your marking is. This gives an indication of the extent to which candidates will receive a reliable outcome for their work from you.

Another purpose of the meeting is to look at actual candidate work and provide any clarification for examiners, such as extra examples of acceptable responses in marking notes.

It is important that you read carefully any markschemes or extra notes that are sent to you before you start marking the practice scripts. This practice stage of the process is your opportunity to gain an understanding of the required marking standard for the session. While marking the practice scripts you can check your marks against the definitive marks and read any comments or annotations written on the scripts by the principal examiner. You may also contact your team leader with any queries you might have.

How the quality model affects you

When you finish marking the practice scripts, you should be confident that you understand the markscheme/criteria/marketing notes and how they should be applied to the current session’s examination. You should have a clear understanding of how and why the definitive marks were awarded to the practice scripts. At this point you can begin the qualification process.

The first qualification stage is your opportunity to show that you can mark to the required standard; for this reason, your team leader is not able to help you. Both the total marks and the marks you award to each question item are compared with the definitive marks. To progress to live marking, you must mark all work within a pre-determined tolerance. A tolerance is applied because it is recognized that there can be legitimate

differences in the marks awarded by different examiners to the same piece of work. It is vitally important that such differences are minimized as much as possible so that there is minimal impact on candidate outcomes. If your marks are outside of the tolerance on one or more scripts, you will receive feedback from a team leader to help you understand how marks should have been awarded.

When the team leader is confident that you understand how marks should have been awarded, he or she will give you access to a second qualification set. You can start live marking if you mark all of these within tolerance. If you pass qualification you have shown that you are able to apply the markscheme accurately and mark in line with the principal examiner meaning that candidates will receive fair outcomes; this is why you are deemed ready to progress onto live marking. However, you should review any differences between your marks and the definitive marks in order to refine your marking and increase the reliability of outcomes for the candidates.

To check that you are continuing to apply the markscheme accurately and in line with the required marking standard, you will receive seeds at regular intervals throughout your live marking. If you mark a seed within tolerance you can view the definitive marks and comments to identify any areas where your marking differed from the definitive marks; this helps you to refine your marking to ensure you mark closely in line with the principal examiner which ensures reliable outcomes for the candidates. If you mark a seed outside of tolerance you must view the definitive marks and annotations in order to understand why your marking was out of line before you can continue marking. If you continue to mark seeds outside of tolerance you will be suspended from marking. This is to prevent candidates from receiving marks that are inaccurate and unreliable.

QIGing – what does it mean?

In a number of components the examination is split into smaller sections to mark. Each section is called a question item group (QIG). Splitting the exam in this way means that you use a smaller part of the markscheme, which reduces the cognitive load. You may find different QIGs easier or more enjoyable to mark. An examination can also be split by option or topic, allowing you to mark options or topics in which you feel more confident.

Because QIGing involves marking on-screen, the quality model used is seeding. You will need to complete practice and qualification stages for each QIG you wish to mark. During live marking your progress is monitored by seeds.

Moderating internal assessment

Some skills, for example those that are assessed over time, are more suited to being assessed by a teacher than an examiner. Internally assessed components are externally moderated by examiners using samples of candidates' work. Examiners who moderate internal assessment are often referred to as "moderators" to highlight that they are moderating marked work rather than examining unmarked work. The samples are selected by IBIS and should be representative of the range of marking within each school.

All internally assessed components are marked by applying assessment criteria or markbands. In the majority of cases the teacher has access to considerably more information about the context and process underlying a candidate's work than the examiner. Because of this, when moderating you are asked to judge whether the teacher's marking seems appropriate rather than to simply re-mark the work and disregard the marks awarded by the teacher. A teacher's marks should be altered only when you are sure they are inappropriate.

Your marks and the marks awarded by the teacher are compared using linear regression, and an appropriate adjustment is applied to each school to bring the teacher's marking in line with yours. This adjustment is referred to as a moderation factor. You might also need to send a sample to your team leader. If so, the marks will be compared in the same way to see if a further adjustment is needed to bring your marking in line with your team leader's standard. Your team leader, in turn, will send a sample of his or her own marking to the principal examiner. This moderation process means you will be unaware of any adjustment being made to your marking when you write each school's internal assessment (IA) feedback, so you should not make any comments about the differences between your marks and the teacher's marks.

Dynamic sampling for moderating on screen

An alternative quality model, which combines features of both moderation and seeding, is used for internally assessed work that is e-marked. Initially you will moderate the work of three candidates from each school's sample. If the differences between the teacher's marks and your marks are within a pre-determined tolerance, it is deemed that the teacher's marking is appropriate and no adjustment will be applied to his or her marks. If one or more is outside of tolerance, you will be given the rest of the sample to moderate and an appropriate moderation factor will be calculated and applied to the teacher's marks.

In the dynamic sampling model, you need to qualify for marking in a similar way to the qualification process for externally assessed e-marked components. You will be monitored using seeds instead of sending a sample of your marking to a more senior examiner. Because you mark using the same standard as the principal examiner, no moderation adjustment is applied to your marking. This enables you to write detailed and meaningful comments in the IA feedback, including comments about the differences between your marks and the teacher's marks.

Moderating external assessment

For the few remaining externally assessed components that are not currently e-marked the moderation quality model is used. The process is similar to the moderation of internal assessment described above.

The process of moderation relies on your marking remaining consistent throughout the marking period, so unlike in the seeding model you do not receive feedback on your marking until all marking (including enquiry upon results marking) is complete. If your sample shows that your marking is inconsistent or that you have not applied the criteria/markbands accurately, a moderation adjustment cannot be applied and your work will be re-marked by another examiner to ensure that candidates receive fair outcomes.

Grade award – how are boundaries set?

In the IB, *marking* and *grading* student work are two different actions.

In marking, you give a candidate credit for the work he or she has produced against a markscheme or similar framework. This shows how much of the assessment task the candidate got right. The mark itself has no other meaning.

In deciding a grade, a judgment is made on the quality of the work against a defined standard. It will take into account the difficulty of the task as well as the number of marks given. Grades have a particular meaning and relevance and are intended to be comparable with grades in other subjects.

Consider: *Marks* are how far a candidate has walked, but *grades* take into account how steep the path was.

At the end of the marking period a grade award meeting is held in every subject. The purpose of the meeting is for the senior examiners, along with the subject manager, to review each component and set grade boundaries. Prior to the meeting the senior examiners and subject manager have access to teachers' comments submitted on G2 forms about each component. These comments from teachers are important background information for the meeting. You also help to provide important background information to the meeting by completing your examiner report, which comments on the performance of candidates in your allocation. In addition to their use at grade award, examiners' reports also contribute to the final subject report that is published to schools.

The grade boundaries for internally assessed and non-examination (coursework) components are set at the beginning of the course and are not revised each session because the task does not change. The boundaries for these components are instead reviewed to check that they continue to be appropriate. However, new boundaries are set for each examination component in every session, due to the fact that a new examination is set for every session. Both the senior examiners' professional judgment and statistical information are used to determine where the boundaries should be. If a significant change in the performance of candidates within a component or subject is identified and can be accounted for, for example, if the current session's examination is known to be more difficult than previous sessions' examinations, the grade boundaries will be changed accordingly. It is the duty of the grade award team to ensure that the same grade standard set in previous sessions is carried forward.

“At risk” re-marking

After grade boundaries have been set, some extra checks on the marking are carried out to ensure the grades issued to candidates are fair and correct. One of these extra checks is “at risk” re-marking. These are candidates who are identified as being at risk of their subject grade being incorrectly low. Some “at risk” re-marking concentrates on candidates who are two or more grades below predicted and require one or two scaled marks to obtain a higher grade. Other “at risk” re-marking concentrates on the work of examiners whose marking has been identified as potentially problematic. “At risk” re-marking is carried out not to identify whether extra marks can be found, but instead as a check that the current mark is appropriate to ensure candidates receive the correct grades.

Enquiry upon results – is your marking more generous?

If schools feel that the marks awarded to their candidates are incorrect, they may request enquiries upon results (EURs).

A category 1 EUR is a re-mark of a candidate’s externally assessed components for one subject. The re-marking is normally carried out by senior examiners. Similar to “at risk” re-marking, when re-marking a candidate’s work for an EUR, an examiner is not looking to award the candidate extra marks but is checking the previous marking to correct any errors identified. If no errors are identified, then the current mark is appropriate. If the examiner identifies errors in the previous marking, they are asked to correct the marking and make a comment to explain the change in marks or give details of the error made. This comment will be referred to by the IB when investigating the original marking that was deemed erroneous by the re-marker.

Schools may also request the return of candidate work through a category 2 EUR. Although you do not have to write comments on the scripts as feedback for teachers, it is important that any comments you do write are appropriate and relate to the assessment criteria/markbands/markscheme.

A category 3 EUR is a re-moderation of a school’s internal assessment sample. This EUR is only available to schools where a significant reduction was applied to the teacher’s marks. The re-moderation is normally carried out by senior examiners. When re-moderating a school’s work, it is important that the examiner uses the same standard that they used in the session. If an examiner makes changes to the original examiner’s marks, it is helpful to write a comment in the sample screen. This comment will be referred to by the IB when investigating the original moderation that was deemed inappropriate by the re-moderator.