

International Baccalaureate Diploma Programme

A guide to assessment

The IB Diploma Programme (DP) is a rigorous, academically challenging and balanced programme of education designed to prepare students aged 16 to 19 for success at university and in life. The DP aims to encourage students to be knowledgeable, inquiring, caring and compassionate, and to develop intercultural understanding, open-mindedness and the attitudes necessary to respect and evaluate a range of viewpoints.

The aim of the DP, like all IB programmes, is to develop internationally minded people who, recognizing their common humanity and shared guardianship of the planet, help to create a better and more peaceful world. The DP provides the opportunity to develop both disciplinary and interdisciplinary knowledge that meets the rigorous standards set by institutions of higher learning around the world.

To ensure both the breadth and depth of knowledge and understanding, students must choose at least one subject from five groups: 1) their best language, 2) additional language(s), 3) humanities and social sciences, 4) experimental sciences, and 5) mathematics, and either an arts subject, or a second subject from the groups 1 to 5. Of these six subjects, three are taken at the higher level (HL) and three are taken at the standard level (SL). In addition, three core elements—the extended essay, theory of knowledge and creativity, action, service—are compulsory and central to the philosophy of the programme.



Overview

The DP utilizes both internally and externally-assessed components to assess student performance. Because of their objectivity and reliability, written examinations at the end of the DP form the basis of the assessment for most courses. Externally assessed coursework completed by students over an extended period under authenticated teacher supervision forms part of the assessment for several programme areas, including theory of knowledge (TOK) essays and the extended essay (EE). In most subjects, students also complete in-school assessment tasks, which are either externally assessed or marked by teachers and then moderated by the IB. By its nature, DP assessment is summative, designed to record student achievement towards the end of the course of study. However, many of the assessment instruments, particularly internal assessment tasks, are also used formatively throughout the teaching and learning process.

DP scores

IB Diploma Programme assessment is criterion-related, rather than measured against the performance of other exam takers. Student performance is measured, using a variety of different methods, against the characteristics of the work expected of each grade level (ie grade descriptors), reflecting the aims and objectives of each subject.

Students receive grades ranging from 7 (highest) to 1 (lowest) for each DP course attempted. A student's final diploma score is made up of the combined scores for each subject. The diploma is awarded to students who gain at least 24 points, subject to certain minimum levels of performance—including successful completion of the 3 elements of the core. TOK and EE are awarded individual grades and collectively can contribute up to three additional points towards the overall diploma score. CAS does not contribute to the points total, but authenticated participation is a requirement for the award of the diploma.

A bilingual diploma is awarded to a candidate who receives a grade of 3 or higher in two languages selected from studies in language and literature. It can also be achieved by a candidate who gains a grade of 3 or higher in studies in language and literature and a grade of 3 or higher in an individuals and societies or science subject completed in a different language.

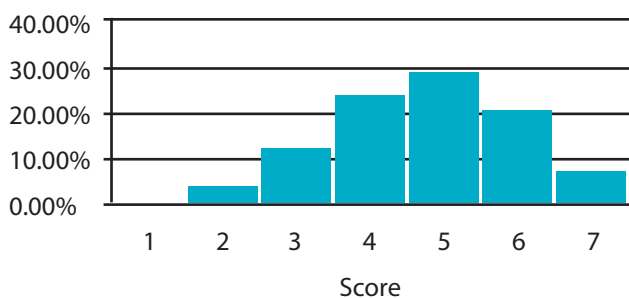
Higher level versus standard level courses

Awarding the same number of points for both HL and SL courses reflects the IB philosophy of the importance of achievement across a broad range of academic disciplines. HL and SL courses differ in scope but are assessed against the same grade descriptors, with HL candidates expected to demonstrate the various elements of the grade descriptors across a greater body of knowledge, understanding and skills.

Student performance

The maximum possible diploma points total, 45 (6 courses x 7 points + 3 total points for the EE and TOK), is achieved by less than 1% of candidates. About 5% of candidates gain more than 40 points. The average score is around 30 points. Around 80% of DP students achieve the diploma each examination session. The pass rate has remained statistically stable over the years, pointing to the consistency of Diploma Programme assessment practices.

Percent distribution of DP exam scores, 2013



Aims and approaches

Assessment of the DP is based on the following aims:

1. to support the curricular and philosophical goals of the programme through the encouragement of good classroom practice and appropriate student learning
2. to ensure that assessment results have a sufficiently high level of reliability, appropriate to a high-stakes university entrance qualification
3. to reflect the international-mindedness of the programme, avoid cultural bias, and make appropriate allowance for students working in their second language
4. to emphasize higher-order cognitive skills (synthesis, reflection, evaluation, critical thinking)
5. to include a suitable range of tasks and instruments/components that ensure all objectives for the subject are assessed
6. to determine student achievement and subject grades through the professional judgment of experienced senior examiners supported by statistical information.

1. Supporting curriculum goals and the programme philosophy

The DP curriculum structure defines the framework in which assessment operates. DP assessments support and encourage the teaching and learning intended by the DP curriculum, as well as providing evidence for the certification of achievement for university admission. Supporting appropriate student learning is the highest priority of the assessments. The strong impact of high-stakes assessment on teaching and learning is used advantageously via valid assessment instruments that encourage good pedagogy and constructive student involvement in their own learning. The DP aim of encouraging students to be “inquiring, knowledgeable and caring” and become “active, compassionate and lifelong learners” (as stated in the IB mission statement) is also reflected in the assessments.

2. The reliability of results

The IB uses about 8,500 examiners worldwide, predominantly experienced DP teachers, who ensure that student work is assessed fairly and consistently. Each subject has a chief examiner, usually an academic from higher education, and a group of senior examiners who prepare examination questions, set the standard for marking and determine the marks needed for the award of each subject grade.

The reliability of results must be a priority for any high-stakes assessment system. To reduce threats to reliability, DP assessments consist of a variety of tasks undertaken in different contexts and on different occasions. Emphasis is given to determining grades that consistently represent the same standard of achievement.

Marker reliability is also essential. As the IB now e-marks virtually all externally assessed work, quality control systems, including detailed markschemes, assessment criteria and sophisticated monitoring procedures ensure that individual examiners mark as closely to the set standards as possible.

Great care is also taken to ensure **grading reliability** in determining grade boundaries through the application of consistent standards supported by statistical data.

There are several ways of ensuring high **reliability in the grading process**:

- using only examiners with appropriate qualifications and experience, who can mark consistently and objectively
- providing examiners with comprehensive instruction and training on how to mark via analytic markschemes and assessment criteria; analytic markschemes give specific instruction regarding how to mark different parts of the examination questions, and assessment criteria are applied when an assessment task is too open-ended to permit detailing all possible valid responses via markschemes
- using standardization scripts to ensure that examiners have the ability to apply the markscheme according to set standards
- randomly assigning examiners “seed” scripts marked by the senior examining team, thus alerting the IB if examiners mark outside the limits of tolerance
- after marking takes place:
 - investigations are carried out on “at risk” candidates who meet the following criteria: either their final grade is two or more grades worse than predicted and are they within two percentage points of getting a better subject grade or their final grade is one grade worse than predicted, they are within two percentage points of getting a better subject grade and at least one component was marked by an examiner who marked seeds at or beyond the acceptable boundaries of tolerance
 - evidence may emerge of unsatisfactory or inconsistent examiner marking while re-marking “enquiries upon results”.

3. Reflecting international-mindedness and avoiding cultural bias

Assessments are carefully developed and include a wide range of formats to minimize bias toward particular cultural groups, and so that students from diverse cultural backgrounds encounter both familiar and unfamiliar contexts and tasks. DP courses are tolerant of cultural variations as well as encouraging of cultural tolerance. Beyond its academic aims, the IB intends that DP students should develop as “caring young people who help to create a better and more peaceful world through intercultural understanding and respect . . . who understand that other people, with their differences, can also be right” (as stated in the IB mission statement). Therefore, both the international context and intercultural purpose of teaching must be reflected in the assessment. This is achieved through:

- the study of language
- the significant cross-cultural dimensions included in many DP subjects
- practical interculturalism encouraged in the arts subjects
- possibilities for practical involvement that arise in the internal assessment work for some other subjects
- beyond knowledge and understanding, attitude and action elements reflected in the CAS requirement

4. Emphasizing higher-order cognitive skills

The more valued academic skills of today are those of accessing, ordering, sifting, synthesizing and evaluating information, and creatively constructing rather than reproducing knowledge. Thus, learning to learn becomes more valuable than just learning content. To this end, DP curriculum objectives and assessments give significant prominence to the “higher-order” cognitive skills. Student analysis, synthesis and evaluation skills can only be properly gauged by requiring them to analyse, synthesize and evaluate at some length. Performance assessment is the only realistic means of assessing student achievement in these areas, and because the outcomes of such activity cannot be tightly prescribed, comparatively unstructured and open-ended assessments are used, including substantial tasks requiring students to reflect on their knowledge and construct extended responses to the task set.

5. A suitable range of assessment components

The course objectives define the construct being assessed, provide a framework for course content, influence teaching, determine assessment tasks and instruments, and provide guidance to markers. Because the objectives represent a wide variety of skill types, assessment tasks and components likewise vary considerably within and across subjects. The types of components include: essays, structured problems, short answer questions, responses to data or text, case studies, and some use of multiple choice questions.

As part of the seven-year curriculum review cycle for each subject, each assessment model is revised by a group of experienced teachers, examiners, IB staff and external consultants to ensure:

- the assessment model provides an appropriate, valid and authentic assessment of the syllabus

- the academic standards and comparability of different courses
- minimal overlap of subject content or objectives
- courses offered complement each other
- the overall assessment load on students, teachers and the IB is manageable
- unnecessary duplication of assessment is eliminated.

An assessment instrument/component is made up of one or more tasks. Examination papers, a portfolio of work, a project, or a research assignment are examples of components. While the majority of components are examination papers containing a range of question types, or tasks, a wide variety of assessments are used to:

- fit the requirements of the subject
- help reduce the potential for inequity in assessment
- ensure that student achievement with regards to all the objectives for a subject is adequately represented.

Internal assessments

Internal or teacher assessments are used for most courses. This may include: oral work, fieldwork, laboratory work, or artistic performances. These normally contribute between 20% and 30% of the subject assessment, but can account for as much as 50% in some courses. Internal assessments allow students to provide evidence of achievement against objectives that do not lend themselves to external examination. Such work can be very flexible in the choice of topic, making internal assessment a valuable addition to students’ education and improving the validity of the assessment process and learning experience as a whole. To ensure the marking reliability of internally assessed work, every school has a sample of their marking re-marked by a moderator. Statistical comparisons and linear regression techniques are used to determine the degree to which the original teacher’s marks need adjusting to bring them in line with the set standards.

6. The role of professional judgment

The complex, higher-order cognitive skills that form the focus of DP assessment do not lend themselves readily to atomized, deconstructed marking. Student responses to many assessment tasks can be highly varied, with many equally valid and correct forms of response. Thus, much depends on the professional judgment of markers, and particularly on the professional expertise of the senior examiners. Considerable effort is expended on providing instruction, guidance and support material to all markers, and on developing quality control systems to check marking standards. Grade awarding in the IB is a balance between statistical data and the judgment of senior examiners of IB student performance against the expected standards embodied in the grade descriptors.

The process of preparing, releasing and reviewing assessments

Examination paper preparation

DP examination papers, developed by diverse teams of experts, have a high level of construct validity, an appropriate level of intellectual demand, and are as free as possible from bias. After a complete set of papers is prepared, they are reviewed multiple times by external experts to determine whether the course content and objectives are covered. Markschemes and explanatory marking notes are developed using the same thorough processes (given their crucial importance to the integrity of the overall assessment).

Setting grade boundaries and distribution

The setting of grade boundaries incorporates the experienced judgment of senior examiners, statistical comparisons and the expectations of experienced teachers. The grade boundaries for each examination paper that have the greatest impact on candidates' progression into higher education (4, 7 and 3) are determined judgmentally in that order by a review of the quality of candidate work against grade descriptors. Assessment component boundaries are compiled into subject grade boundaries, leading to the subject grade distribution, which is then checked against several indicators. This approach allows for more transparency than more complex models of weighting and scaling.

Publication of results

Soon after the official results are released, the final candidate marks for each assessment component are made available, as well as the examination papers and their associated markschemes. Subject reports are written by the senior examining team, covering all general aspects of candidate performance on each component, outlining where candidates performed well, where they performed less well, and making recommendations for improving the preparation of future candidates.

Feedback and enquiries upon results

After the release of results, schools have a number of avenues to provide feedback and challenge assessment outcomes. The "enquiry upon results" service allows schools to request a re-mark of a candidate's work if they feel the result is not a fair reflection of their performance. Subject grades may be raised or lowered. Schools may also request the return of copies of the externally marked work for a given component, allowing teachers to see how a piece of work has been marked. Re-moderation of an internally assessed coursework sample is also available in cases where a teacher's marks were reduced by an average of 15% or more of the component's maximum mark through moderation.

About the IB: For more than 40 years the IB has built a reputation for high-quality, challenging programmes of education that develop internationally minded young people who are well prepared for the challenges of life in the 21st century and able to contribute to creating a better, more peaceful world.

For further information on the IB Diploma Programme, visit www.ibo.org/diploma. Complete subject guides can be accessed through the IB online curriculum centre, the IB university and government official system, or purchased through the IB store: store.ibo.org.

To learn more about how the IB Diploma Programme prepares students for success at university, visit www.ibo.org/recognition or email recognition@ibo.org.



International Baccalaureate®
Baccalauréat International
Bachillerato Internacional