

Assessment in the Diploma Programme

General

Assessment is an integral part of teaching and learning. The most important aims of assessment in the Diploma Programme are that it should support curricular goals and encourage appropriate student learning. Both external and internal assessments are used in the Diploma Programme. IB examiners mark work produced for external assessment, while work produced for internal assessment is marked by teachers and externally moderated by the IB.

There are two types of assessment identified by the IB.

- Formative assessment informs both teaching and learning. It is concerned with providing accurate and helpful feedback to students and teachers on the kind of learning taking place and the nature of students' strengths and weaknesses in order to help develop students' understanding and capabilities. Formative assessment can also help to improve teaching quality, as it can provide information to monitor progress towards meeting the course aims and objectives.
- Summative assessment gives an overview of previous learning and is concerned with measuring student achievement.

The Diploma Programme primarily focuses on summative assessment designed to record student achievement at, or towards the end of, the course of study. However, many of the assessment instruments can also be used formatively during the course of teaching and learning, and teachers are encouraged to do this. A comprehensive assessment plan is viewed as being integral with teaching, learning and course organization. For further information, see the IB *Programme standards and practices* document.

The approach to assessment used by the IB is criterion-related, not norm-referenced. This approach to assessment judges students' work by their performance in relation to identified levels of attainment, and not in relation to the work of other students. For further information on assessment within the Diploma Programme please refer to the publication *Diploma Programme assessment: Principles and practice*.

To support teachers in the planning, delivery and assessment of the Diploma Programme courses, a variety of resources can be found on the OCC or purchased from the IB store (<http://store.ibo.org>). Additional publications such as specimen papers and markschemes, Teacher support material, subject reports and grade descriptors can also be found on the online curriculum centre (OCC). Past examination papers as well as markschemes can be purchased from the IB store.

Methods of assessment

The IB uses several methods to assess work produced by students.

Assessment criteria

Assessment criteria are used when the assessment task is open-ended. Each criterion concentrates on a particular skill that students are expected to demonstrate. An assessment objective describes what students should be able to do, and assessment criteria describe how well they should be able to do it. Using assessment criteria allows discrimination between different answers and encourages a variety of responses. Each criterion comprises a set of hierarchically ordered level descriptors. Each level descriptor is worth one or more marks. Each criterion is applied independently using a best-fit model. The maximum marks for each criterion may differ according to the criterion's importance. The marks awarded for each criterion are added together to give the total mark for the piece of work.

Markbands

Markbands are a comprehensive statement of expected performance against which responses are judged. They represent a single holistic criterion divided into level descriptors. Each level descriptor corresponds to a range of marks to differentiate student performance. A best-fit approach is used to ascertain which particular mark to use from the possible range for each level descriptor.

Analytic markschemes

Analytic markschemes are prepared for those examination questions that expect a particular kind of response and/or a given final answer from students. They give detailed instructions to examiners on how to break down the total mark for each question for different parts of the response.

Marking notes

For some assessment components marked using assessment criteria, marking notes are provided. Marking notes give guidance on how to apply assessment criteria to the particular requirements of a question.

Assessment of the extended essay

Assessment of the extended essay is a combination of formative assessment (the *Reflections on planning and progress form*) and summative assessment (the extended essay itself).

Generic assessment criteria are used with subject-specific interpretations.

Inclusive assessment arrangements

Inclusive assessment arrangements are available for students with assessment access requirements. These arrangements enable students with diverse needs to access the examinations and demonstrate their knowledge and understanding of the constructs being assessed.

The IB document *Candidates with assessment access requirements* provides details on all the inclusive assessment arrangements available to students with learning support requirements. The IB document *Learning diversity within the International Baccalaureate programmes: Special educational needs within the International Baccalaureate programmes* outlines the position of the IB with regard to students with diverse learning needs in the IB programmes.

For students affected by adverse circumstances, the IB documents *General regulations: Diploma Programme* and the *Handbook of procedures for the Diploma Programme* provide details on access consideration.

Responsibilities of the school

The school is required to ensure that equal access arrangements and reasonable adjustments are provided to students with learning support requirements that are in line with the IB documents *Candidates with assessment access requirements* and *Learning diversity within the International Baccalaureate programmes: Special educational needs within the International Baccalaureate programmes*.

Overview of the assessment criteria for the extended essay

Overview

Criterion A: focus and method	Criterion B: knowledge and understanding	Criterion C: critical thinking	Criterion D: presentation	Criterion E: engagement
<ul style="list-style-type: none">• Topic• Research question• Methodology	<ul style="list-style-type: none">• Context• Subject-specific terminology• and concepts	<ul style="list-style-type: none">• Research• Analysis• Discussion and evaluation	<ul style="list-style-type: none">• Structure• Layout	<ul style="list-style-type: none">• Process• Research focus
Marks	Marks	Marks	Marks	Marks
6	6	12	4	6

Total marks available: 34

Best-fit approach and markbands

Assessment criteria must be used in conjunction with the relevant specific-subject interpretations as these interpretations articulate how the generic assessment criteria are understood and applied to each subject.

Levels of performance are described using multiple indicators per level. In many cases the indicators occur together throughout the essay, but not always. Also, not all indicators are always present. This means that a student can demonstrate performances that fit into different levels. To accommodate this, the IB assessment models use markbands and advise examiners and teachers to use a **best-fit approach** in deciding the appropriate mark for a particular criterion. From various assessment trials we know that introducing markbands and using the best-fit model is not always self-evident, and guidance is needed to help with their application. While the extended essay is an externally assessed component of the DP, supervisors are required to submit a predicated grade and understanding the way in which the criteria are applied by examiners will assist with the guidance given to students. The following explains how markbands are used by examiners. The aim is to find the descriptor that conveys most accurately the level attained by the student's work, using the best-fit approach. A best-fit approach means that compensation will be made when a piece of work matches different aspects of a markband at different levels. The mark awarded will be one that most fairly reflects the balance of achievement against the markband. It is not necessary for every indicator of a level descriptor to be met for that mark to be rewarded. (For example, if student work matches two of the three requirements within a markband but one is seriously lacking, the student should be awarded for the strands that have been met well, but the mark awarded should be at the lower end of the markband to compensate for what is lacking in one strand. If the level of student work spans multiple markbands, compensation depends on the performance in the higher order skills of evaluation (AO3), discussion (AO3) and analysis (AO2) (see the example below). The assessment objective levels for a given subject can be found at the back of that subject's DP subject guide.

Criterion C: Critical thinking

Research	Excellent (10–12)	Research	Good (7–9)
Analysis	Good (7–9)	Analysis	Good (7–9)
Discussion/evaluation	Adequate (4–6)	Discussion/evaluation	Adequate (4–6)
Mark awarded	8/9 (The 7–9 markband is appropriate because communication of research is a lower order skill compared to analysis and evaluation.)	Mark awarded	7 (The bottom end of the 7–9 markband is appropriate since the achievement level is lower for the higher order skill of discussion/evaluation.)

- When assessing a student's work, in light of the IB approach to positive marking, examiners will read the level descriptors from the highest markband down until they reach a descriptor that most appropriately describes the level of the work being assessed.

- If a piece of work seems to fall between two descriptors, both descriptors will be read again and the one that more appropriately describes the student's work will be chosen. In relation to criterion C, examiners will bear in mind the higher order skills being assessed.
- There are a number of marks available within a level; examiners will award the upper marks if the student's work demonstrates the qualities described to a greater extent. Examiners will award the lower marks if the student's work demonstrates the qualities described to a lesser extent.
- The highest level descriptors do not imply faultless performance and should be achievable by a student. Examiners will not hesitate to use the extremes if they are appropriate descriptions of the work being assessed.

The assessment criteria

Criterion A: Focus and method

This criterion focuses on the topic, the research question and the methodology. It assesses the explanation of the focus of the research (this includes the topic and the research question), how the research will be undertaken, and how the focus is maintained throughout the essay.

Level	Descriptor of strands and indicators
0	The work does not reach a standard outlined by the descriptors below.
1–2	The topic is communicated unclearly and incompletely. <ul style="list-style-type: none">Identification and explanation of the topic is limited; the purpose and focus of the research is unclear, or does not lend itself to a systematic investigation in the subject for which it is registered. The research question is stated but not clearly expressed or too broad. <ul style="list-style-type: none">The research question is too broad in scope to be treated effectively within the word limit and requirements of the task, or does not lend itself to a systematic investigation in the subject for which it is registered.The intent of the research question is understood but has not been clearly expressed and/or the discussion of the essay is not focused on the research question. Methodology of the research is limited. <ul style="list-style-type: none">The source(s) and/or method(s) to be used are limited in range given the topic and research question.There is limited evidence that their selection was informed.
3–4	The topic is communicated. <ul style="list-style-type: none">Identification and explanation of the research topic is communicated; the purpose and focus of the research is adequately clear, but only partially appropriate. The research question is clearly stated but only partially focused. <ul style="list-style-type: none">The research question is clear but the discussion in the essay is only partially focused and connected to the research question. Methodology of the research is mostly complete. <ul style="list-style-type: none">Source(s) and/or method(s) to be used are generally relevant and appropriate given the topic and research question.There is some evidence that their selection(s) was informed. If the topic or research question is deemed inappropriate for the subject in which the essay is registered no more than four marks can be awarded for this criterion.
5–6	The topic is communicated accurately and effectively.

Level	Descriptor of strands and indicators
	<ul style="list-style-type: none"> Identification and explanation of the research topic is effectively communicated; the purpose and focus of the research is clear and appropriate. <p>The research question is clearly stated and focused.</p> <ul style="list-style-type: none"> The research question is clear and addresses an issue of research that is appropriately connected to the discussion in the essay. <p>Methodology of the research is complete.</p> <ul style="list-style-type: none"> An appropriate range of relevant source(s) and/or method(s) have been applied in relation to the topic and research question. There is evidence of effective and informed selection of sources and/or methods.

Criterion B: Knowledge and understanding

This criterion assesses the extent to which the research relates to the subject area/discipline used to explore the research question, or in the case of the world studies extended essay, the issue addressed and the two disciplinary perspectives applied, and additionally the way in which this knowledge and understanding is demonstrated through the use of appropriate terminology and concepts.

Level	Descriptor of strands and indicators
0	The work does not reach a standard outlined by the descriptors below.
1–2	<p>Knowledge and understanding is limited.</p> <ul style="list-style-type: none"> The selection of source material has limited relevance and is only partially appropriate to the research question. Knowledge of the topic/discipline(s)/issue is anecdotal, unstructured and mostly descriptive with sources not effectively being used. <p>Use of terminology and concepts is unclear and limited.</p> <ul style="list-style-type: none"> Subject-specific terminology and/or concepts are either missing or inaccurate, demonstrating limited knowledge and understanding.
3–4	<p>Knowledge and understanding is good.</p> <ul style="list-style-type: none"> The selection of source material is mostly relevant and appropriate to the research question. Knowledge of the topic/discipline(s)/issue is clear; there is an understanding of the sources used but their application is only partially effective. <p>Use of terminology and concepts is adequate.</p> <ul style="list-style-type: none"> The use of subject-specific terminology and concepts is mostly accurate, demonstrating an appropriate level of knowledge and understanding. <p>If the topic or research question is deemed inappropriate for the subject in which the essay is registered no more than four marks can be awarded for this criterion.</p>

Level	Descriptor of strands and indicators
5–6	<p>Knowledge and understanding is excellent.</p> <ul style="list-style-type: none"> The selection of source materials is clearly relevant and appropriate to the research question. Knowledge of the topic/discipline(s)/issue is clear and coherent and sources are used effectively and with understanding. <p>Use of terminology and concepts is good.</p> <ul style="list-style-type: none"> The use of subject-specific terminology and concepts is accurate and consistent, demonstrating effective knowledge and understanding.

Criterion C: Critical thinking

This criterion assesses the extent to which critical-thinking skills have been used to analyse and evaluate the research undertaken.

Level	Descriptor of strands and indicators
0	The work does not reach a standard outlined by the descriptors below.
1–3	<p>The research is limited.</p> <ul style="list-style-type: none"> The research presented is limited and its application is not clearly relevant to the RQ. <p>Analysis is limited.</p> <ul style="list-style-type: none"> There is limited analysis. Where there are conclusions to individual points of analysis these are limited and not consistent with the evidence. <p>Discussion/evaluation is limited.</p> <ul style="list-style-type: none"> An argument is outlined but this is limited, incomplete, descriptive or narrative in nature. The construction of an argument is unclear and/or incoherent in structure hindering understanding. Where there is a final conclusion, it is limited and not consistent with the arguments/evidence presented. There is an attempt to evaluate the research, but this is superficial. <p>If the topic or research question is deemed inappropriate for the subject in which the essay is registered no more than three marks can be awarded for this criterion.</p>
4–6	<p>The research is adequate.</p> <ul style="list-style-type: none"> Some research presented is appropriate and its application is partially relevant to the Research question. <p>Analysis is adequate.</p>

Level	Descriptor of strands and indicators
	<ul style="list-style-type: none"> There is analysis but this is only partially relevant to the research question; the inclusion of irrelevant research detracts from the quality of the argument. Any conclusions to individual points of analysis are only partially supported by the evidence. <p>Discussion/evaluation is adequate.</p> <ul style="list-style-type: none"> An argument explains the research but the reasoning contains inconsistencies. The argument may lack clarity and coherence but this does not significantly hinder understanding. Where there is a final or summative conclusion, this is only partially consistent with the arguments/evidence presented. The research has been evaluated but not critically.
7–9	<p>The research is good.</p> <ul style="list-style-type: none"> The majority of the research is appropriate and its application is clearly relevant to the research question. <p>Analysis is good.</p> <ul style="list-style-type: none"> The research is analysed in a way that is clearly relevant to the research question; the inclusion of less relevant research rarely detracts from the quality of the overall analysis. Conclusions to individual points of analysis are supported by the evidence but there are some minor inconsistencies. <p>Discussion/evaluation is good.</p> <ul style="list-style-type: none"> An effective reasoned argument is developed from the research, with a conclusion supported by the evidence presented. This reasoned argument is clearly structured and coherent and supported by a final or summative conclusion; minor inconsistencies may hinder the strength of the overall argument. The research has been evaluated, and this is partially critical.
10–12	<p>The research is excellent.</p> <ul style="list-style-type: none"> The research is appropriate to the research question and its application is consistently relevant. <p>Analysis is excellent.</p> <ul style="list-style-type: none"> The research is analysed effectively and clearly focused on the research question; the inclusion of less relevant research does not significantly detract from the quality of the overall analysis. Conclusions to individual points of analysis are effectively supported by the evidence. <p>Discussion/evaluation is excellent.</p>

Level	Descriptor of strands and indicators
	<ul style="list-style-type: none"> An effective and focused reasoned argument is developed from the research with a conclusion reflective of the evidence presented. This reasoned argument is well structured and coherent; any minor inconsistencies do not hinder the strength of the overall argument or the final or summative conclusion. The research has been critically evaluated.

Criterion D: Presentation

This criterion assesses the extent to which the presentation follows the standard format expected for academic writing and the extent to which this aids effective communication.

Level	Descriptor of strands and indicators
0	The work does not reach a standard outlined by the descriptors below.
1–2	<p>Presentation is acceptable.</p> <ul style="list-style-type: none"> The structure of the essay is generally appropriate in terms of the expected conventions for the topic, argument and subject in which the essay is registered. Some layout considerations may be missing or applied incorrectly. Weaknesses in the structure and/or layout do not significantly impact the reading, understanding or evaluation of the extended essay.
3–4	<p>Presentation is good.</p> <ul style="list-style-type: none"> The structure of the essay clearly is appropriate in terms of the expected conventions for the topic, the argument and subject in which the essay is registered. Layout considerations are present and applied correctly. The structure and layout support the reading, understanding and evaluation of the extended essay.

Criterion E: Engagement

This criterion assesses the student's engagement with their research focus and the research process. It will be applied by the examiner at the end of the assessment of the essay, and is based solely on the candidate's reflections as detailed on the [RPPF](#), with the supervisory comments and extended essay itself as context.

Level	Descriptor of strands and indicators
0	The work does not reach a standard outlined by the descriptors or a RPPF has not been submitted.
1–2	<p>Engagement is limited.</p> <ul style="list-style-type: none"> Reflections on decision-making and planning are mostly descriptive.

Level	Descriptor of strands and indicators
	<ul style="list-style-type: none"> These reflections communicate a limited degree of personal engagement with the research focus and/or research process.
3–4	<p>Engagement is good.</p> <ul style="list-style-type: none"> Reflections on decision-making and planning are analytical and include reference to conceptual understanding and skill development. These reflections communicate a moderate degree of personal engagement with the research focus and process of research, demonstrating some intellectual initiative.
5–6	<p>Engagement is excellent.</p> <ul style="list-style-type: none"> Reflections on decision-making and planning are evaluative and include reference to the student's capacity to consider actions and ideas in response to challenges experienced in the research process. These reflections communicate a high degree of intellectual and personal engagement with the research focus and process of research, demonstrating authenticity, intellectual initiative and/or creative approach in the student voice.

Assessment grade descriptors for the extended essay

Effective May 2018

Grade descriptors

The extended essay is externally assessed, and as such, supervisors are not expected to mark the essays or arrive at a number to translate into a grade. Predicted grades for all subjects should be based on the **qualitative** grade descriptors for the subject in question. These descriptors are what will be used by senior examiners to set the boundaries for the extended essay in May 2018, and so schools are advised to use them in the same way.

Grade A

Demonstrates effective research skills resulting in a well-focused and appropriate research question that can be explored within the scope of the chosen topic; effective engagement with relevant research areas, methods and sources; excellent knowledge and understanding of the topic in the wider context of the relevant discipline; the effective application of source material and correct use of subject-specific terminology and/or concepts further supporting this; consistent and relevant conclusions that are proficiently analysed; sustained reasoned argumentation supported effectively by evidence; critically evaluated research; excellent presentation of the essay, whereby coherence and consistency further supports the reading of the essay; and present and correctly applied structural and layout elements.

Engagement with the process is conceptual and personal, key decision-making during the research process is documented, and personal reflections are evidenced, including those that are forward-thinking.

Grade B

Demonstrates appropriate research skills resulting in a research question that can be explored within the scope of the chosen topic; reasonably effective engagement with relevant research areas, methods and sources; good knowledge and understanding of the topic in the wider context of the relevant discipline; a reasonably effective application of source material and use of subject-specific terminology and/or concepts; consistent conclusions that are accurately analysed; reasoned argumentation often supported by evidence; research that at times evidences critical evaluation; and a clear presentation of all structural and layout elements, which further supports the reading of the essay.

Engagement with the process is generally evidenced by the reflections and key decision-making during the research process is documented.

Grade C

Demonstrates evidence of research undertaken, which has led to a research question that is not necessarily expressed in a way that can be explored within the scope of the chosen topic; partially effective engagement with mostly appropriate research areas, methods and sources—however, there are some discrepancies in those processes, although these do not interfere with the planning and approach; some knowledge and understanding of the topic in the wider context of the discipline, which is mostly relevant; the attempted application of source material and appropriate terminology and/or concepts; an attempted synthesis of research results with partially relevant analysis; conclusions partly supported by the evidence; discussion that is descriptive rather than analytical; attempted evaluation; satisfactory presentation of the essay, with weaknesses that do not hinder the reading

of the essay; and some structural and layout elements that are missing or are incorrectly applied.

Engagement with the process is evidenced but shows mostly factual information, with personal reflection mostly limited to procedural issues.

Grade D

Demonstrates a lack of research, resulting in unsatisfactory focus and a research question that is not answerable within the scope of the chosen topic; at times engagement with appropriate research, methods and sources, but discrepancies in those processes that occasionally interfere with the planning and approach; some relevant knowledge and understanding of the topic in the wider context of the discipline, which are at times irrelevant; the attempted application of source material, but with inaccuracies in the use of, or underuse of, terminology and/or concepts; irrelevant analysis and inconsistent conclusions as a result of a descriptive discussion; a lack of evaluation; presentation of the essay that at times is illogical and hinders the reading; and structural and layout elements that are missing.

Engagement with the process is evidenced but is superficial, with personal reflections that are solely narrative and concerned with procedural elements.

Grade E (failing condition)

Demonstrates an unclear nature of the essay; a generally unsystematic approach and resulting unfocused research question; limited engagement with limited research and sources; generally limited and only partially accurate knowledge and understanding of the topic in the wider context of the relevant discipline; ineffective connections in the application of source material and inaccuracies in the terminology and/or concepts used; a summarizing of results of research with inconsistent analysis; an attempted outline of an argument, but one that is generally descriptive in nature; and a layout that generally lacks or incorrectly applies several layout and structural elements.

Engagement with the process is limited, with limited factual or decision-making information and no personal reflection on the process.