

香港中學文憑考試

Hong Kong Diploma of Secondary Education Examination

香港中學文憑考試  
評級程序與水平參照成績匯報  
Grading Procedures and  
Standards-referenced Reporting  
in the HKDSE Examination



香港考試及評核局  
Hong Kong  
Examinations and  
Assessment Authority

# Table of Contents

## 目錄

1 Introduction . . . . .	2
簡介	
2 The reporting system . . . . .	3
成績匯報制度	
3 Level descriptors . . . . .	5
等級描述	
4 The grading procedures . . . . .	8
評級程序	
5 Setting the standards . . . . .	11
設定水平	
6 Maintaining the standards . . . . .	14
維持水平	
7 Reporting the results . . . . .	17
匯報成績	
Appendix 1 : Group Ability Index . . . . .	18
附錄 1 : 組別能力指數	
Appendix 2 : Equipercentile method . . . . .	19
附錄 2 : 百分位等值法	
Appendix 3 : Latent trait model . . . . .	20
附錄 3 : 潛在特性模型	
Reference materials . . . . .	22
參考資料	

# Introduction

## 簡介

In the Hong Kong Diploma of Secondary Education (HKDSE) Examination, standards-referenced reporting (SRR) is adopted to report candidates' assessment results. What this means is that for a given subject, candidates' levels of performance are reported with reference to a set of explicit and fixed standards of performance.

SRR is not a new approach in public examinations in Hong Kong. It has been in use for the subjects of Chinese Language and English Language in the Hong Kong Certificate of Education (HKCE) Examination since 2007. This step was undertaken in response to demands that results should be reported with reference to standards that are internationally recognised, transparent, explicit, and that remain constant over time.

To cater for the SRR approach, the Hong Kong Examinations and Assessment Authority (HKEAA) has established a grading system which aims to maintain standards between successive years and inter-subject comparability within the same year.

The purpose of this booklet is to explain the rationale, approaches and key features of the grading procedures and the SRR approach. The HKEAA will review the grading mechanism and the SRR approach in the light of experience from the first administration of the HKDSE Examination in 2012. Schools will be informed of any changes with an adequate period of notice.

香港中學文憑考試採用水平參照模式匯報評核結果，各個科目也會參照各自一套明確而固定的標準，來匯報考生表現所屬的等級。

香港的公開考試並不是首次採用水平參照匯報成績，由 2007 年開始，香港中學會考(會考)中國語文和英國語文科，已採用了水平參照模式匯報成績。採用水平參照乃是要回應以下要求：評級所參照的標準須獲國際認可、透明度高、明確而固定。

為了確保年與年之間的水平穩定，並維持同一年內各科成績的可比性，香港考試及評核局(考評局)確立了一套與水平參照模式相配合的評級制度。

本小冊子旨在解釋評級程序與水平參照成績匯報的原理、方法及重點。考評局將在 2012 年首屆香港中學文憑考試完成後，總結有關經驗並檢討評級程序與水平參照模式。如有更改，考評局將於適當時候通知學校。

## 2 The reporting system 成績匯報制度

There are three categories of HKDSE subjects: Categories A, B and C.

### Category A subjects

The HKEAA is responsible for the grading of all Category A subjects. Candidate performance will be reported on the HKDSE certificate at five levels (1–5), with 5 being the highest. The highest-achieving Level 5 candidates will be awarded Level 5\*\* ; the next highest-achieving Level 5 candidates will be awarded Level 5\* and the remaining Level 5 candidates will receive simply Level 5. Performance below the Level 1 standard for a subject will be designated as 'Unclassified'. This will not be reported on the certificate.

### Category B subjects

The assessment of Category B Applied Learning subjects will be undertaken by the organisations providing relevant courses (course providers). After moderation by the HKEAA, the final results will be recorded on the HKDSE certificate. The results will be reported at two levels: 'Attained' and 'Attained with Distinction'. Performance below 'Attained' will be designated as 'Unattained'. This will not be reported on the certificate.

### Category C subjects

Category C subjects comprise six other language subjects at Advanced Subsidiary (AS) level. Question papers will be set and the marking and grading will be conducted by Cambridge International Examinations (CIE). Results will be reported in five grades (A-E) on the HKDSE certificate, with grade E being the lowest and grade A being the highest. Achievements below grade E will be designated as 'Ungraded'. This will not be reported on the certificate.

香港中學文憑考試的科目可分為三類：甲類、乙類及丙類科目。

### 甲類科目

甲類科目的評級由考評局全權負責。在香港中學文憑證書上，考生的表現將以五個等級(1–5級)匯報，第5級為最高等級。在第5級的考生中，成績最優異的可獲評為5\*\*級；成績次佳的可獲評為5\*級；其餘的則評為5級。表現低於第1級的水平將定為「不予評級」，並不會在證書上匯報。

### 乙類科目

乙類應用學習科目的評核由提供相關課程的機構(課程提供者)負責。考生的成績須經考評局審定，並於香港中學文憑證書上匯報。有關科目的成績將以兩個等級匯報，即「達標」和「達標並表現優異」。表現低於「達標」的水平將被定為「不達標」，並不會在證書上匯報。

### 丙類科目

丙類科目包括六個高級補充(高補)程度的其他語言科目。英國劍橋大學國際考試組(CIE)將負責考試的擬題、閱卷及評級等工作。有關科目的成績將以五個等級(A-E)在香港中學文憑證書上匯報。在五個等級中，E級最低，A級最高，而E級以下的成績將定為「不予評級」，並不會在證書上匯報。

## Special arrangements for individual subjects

### Chinese Language and English Language

In addition to the overall subject level, candidates' performance in individual components (e.g. Reading) will also be reported. Performance below the Level 1 standard for individual components will be designated as 'Unclassified'. Candidates whose performance is designated as 'Unclassified' in some components but who achieve Level 1 or above at subject level will have all component and subject levels reported on the HKDSE certificate. However, if the subject level is below Level 1, neither the subject level nor individual component levels will be reported on the certificate.

In English Language, Paper 1 Reading and Paper 3 Listening and Integrated Skills both comprise a compulsory section and a section where candidates can choose between an easier or a more difficult option. Candidates attempting the more difficult option can attain up to Level 5\*\* in the relevant component, while those who attempt the easier option can only attain up to Level 4.

### Mathematics

A candidate may take the Compulsory Part only or take both the Compulsory Part and one of the two modules in the Extended Part. The results of the Compulsory Part and the Extended Part will be reported separately.

### Combined Science

The overall subject level will be reported, as well as the levels for the two chosen components (Biology, Chemistry, or Physics). Candidates whose performance is designated as 'Unclassified' in one component but who achieve Level 1 or above at the subject level will have all component and subject levels reported on the HKDSE certificate. However, if the subject level is below Level 1, neither the subject level nor individual component levels will be reported on the certificate.

## 個別科目的特殊安排

### 中國語文和英國語文

除了科目的整體等級外，兩個語文科目均會匯報每個分部（例如閱讀）的等級。個別分部的成績若低於第1級，將定為「不予評級」。如考生的科目整體成績為第1級或以上，而有個別分部成績為「不予評級」，則所有分部和科目整體成績均會在香港中學文憑證書上匯報；若科目整體成績在第1級以下，則無論分部或整體科目成績，均不會在證書上匯報。

英國語文科的卷一閱讀和卷三聆聽與綜合能力均設有必答部分和選答部分。除必答部分外，考生可在選答部分選擇較易或較難部分作答。如考生選答較難部分，在相關分部最高可獲第5\*\*級，選答較易部分的考生則最高只能獲得第4級。

### 數學

考生可以選擇只應考必修部分，或應考必修部分及延伸部分兩個單元其中之一。必修部分及延伸部分的成績將分別匯報。

### 組合科學

本科的整體等級及兩個選修部分（生物、化學或物理）的等級均會在香港中學文憑證書上匯報。如考生的科目成績在第1級或以上而其中一分部成績為「不予評級」，則所有分部和科目整體成績均會在證書上匯報；若科目成績在第1級以下，則無論分部或整體科目的成績，均不會在證書上匯報。

# 3 Level descriptors

## 等級描述

Under the standards-referenced reporting system, a set of descriptors is written for each subject to describe what a candidate performing at a given level is typically able to do. The principle used in preparing these descriptors is to describe the achievements that candidates are able to demonstrate in positive terms and to avoid describing what candidates cannot achieve.

By their nature, these descriptors represent 'on-average' statements, and may not precisely describe the actual performance of individual candidates. A given candidate may demonstrate, across the examination papers for a single subject, some elements of performance that are characteristic of more than one level. The assessment process that is applied within each subject aggregates and combines such variation to derive a final result that provides a 'best fit' between a given candidate's performance and a level descriptor.

In addition to the subject level descriptors, a set of generic descriptors has been developed to provide an overall brief description of the performance standards of candidates at different levels across subjects (Table 1). The generic descriptors encompass the major learning outcomes that are common across subjects, including knowledge and understanding of the curriculum, the ability to apply concepts and skills, higher order abilities such as interpretation, analysis, synthesis and evaluation, and the ability to communicate.

The descriptors for individual subjects were developed by a working group that comprised members from different sectors, including officers from the Curriculum Development Institute (CDI), school principals, secondary school teachers, and university subject experts. Close reference was made to samples of the actual performances of candidates in previous public examinations and in the pilot tests which used the HKDSE sample papers. The development of level descriptors was completed in 2009 and they were published in the 'Standards-referenced Reporting Information Packages', together with samples of student work that illustrate the relevant standards.

### Use of level descriptors

Level descriptors are important sources of reference for subject experts to make judgments in grading, standards setting and maintenance. For details, please refer to sections 4 to 6.

The level descriptors also facilitate learning and teaching. Students can make use of the descriptors to set their goals and assess their learning progress. Teachers also know explicitly what they have to do to assist students to achieve a higher level of performance.

Moreover, the level descriptors are useful for other users, such as tertiary institutions or employers, in enabling them to better understand what candidates at different levels of performance can do in order for them to make more appropriate selection decisions.

以水平參照匯報成績，各科均附有一套等級描述，以說明該學科達至某等級的考生的典型能力水平。擬訂這些等級描述的原則是正面描述考生能展示的水平，而避免描述考生不能掌握的能力。

這些等級描述只是「一般性」的口述，未必能準確描述個別考生的實際表現。這是由於某個考生可能在同一科目的不同卷別有跨越多於一個等級的表現。在每科的評核過程中，這些不同的表現會被整合，以最配合考生表現的等級來匯報其成績。

除了各科的等級描述外，考評局也制定了一套共通等級描述（表一），這一套描述涵蓋不同科目，簡略說明考生於不同等級表現的水平。共通等級描述是根據各學科共同的主要學習成果而擬訂，當中包括對課程內容的認識和理解，應用概念和技巧的能力，詮釋、分析、綜合和評價等高階能力，以及傳意能力。

個別科目的等級描述是由一個工作小組所擬訂，該小組的成員來自不同界別，當中包括課程發展主任、校長、中學教師，以及大學的學科專家。他們參考過往公開考試考生實際表現的樣本，和用香港中學文憑考試樣本試卷作先導測驗而取得的答卷樣本，以制訂這些等級描述。擬訂等級描述的工作已經在2009年完成，而有關等級描述及學生表現示例亦已輯錄於各科的《水平參照成績匯報資料套》內。

### 等級描述的用途

等級描述是科目專家在評級、設定水平和維持水平等工作的重要參考資料。有關詳情，可參考第4至6章。

同時，等級描述亦可以促進學與教。學生可參考有關描述，用以制定個人的學習目標和評估學習進度；而教師亦可藉此明確瞭解如何協助學生邁向更高水平。

此外，對大專院校或僱主等而言，等級描述讓他們更清楚瞭解不同水平考生的實際能力，從而作出更合適的甄選決定。

## Table I Generic descriptors

### 表一 共通等級描述

#### Level 5 第五級

Candidates at this level typically demonstrate:

- comprehensive knowledge and understanding of the curriculum and the ability to apply the concepts and skills effectively in diverse and complex unfamiliar situations with insight
- ability to analyse, synthesise and evaluate information from a wide variety of sources
- ability to communicate ideas and express views concisely and logically

達到本等級考生的典型表現如下：

- 對課程內容有廣泛的認識和透徹的理解，能把概念和技巧有效地應用到多元和複雜的不熟悉情境，並顯示深入的見解
- 能分析、綜合和評價廣泛的資料
- 能精簡及邏輯地傳達意念和見解

#### Level 4 第四級

Candidates at this level typically demonstrate:

- good knowledge and understanding of the curriculum and the ability to apply the concepts and skills effectively in unfamiliar situations with insight
- ability to analyse, synthesise and interpret information from a variety of sources
- ability to communicate ideas and express views logically

第四級

達到本等級考生的典型表現如下：

- 對課程內容有良好的認識和理解，能把概念和技巧有效地應用到不熟悉的情境，並顯示深入的見解
- 能分析、綜合和詮釋各種資料
- 能邏輯地傳達意念和見解

## Level 3 第三級

Candidates at this level typically demonstrate:

- adequate knowledge and understanding of the curriculum and the ability to apply the concepts and skills appropriately in different familiar situations
- ability to analyse and interpret information from a variety of sources
- ability to communicate ideas and express views appropriately

達到本等級考生的典型表現如下：

- 對課程內容有足夠的認識和理解，能把概念和技巧適當地應用到不同的熟悉情境
- 能分析和詮釋各種資料
- 能恰當地傳達意念和見解

## Level 2 第二級

Candidates at this level typically demonstrate:

- basic knowledge and understanding of the curriculum and the ability to apply the concepts and skills in familiar situations
- ability to identify and interpret information from straightforward sources
- ability to communicate simple ideas in a balanced way

達到本等級考生的典型表現如下：

- 對課程內容有基本的認識和理解，能把概念和技巧應用到熟悉的情境
- 能辨識和詮釋直接的資料
- 能平衡地傳達簡單意念

## Level 1 第一級

Candidates at this level typically demonstrate:

- elementary knowledge and understanding of the curriculum and the ability to apply the concepts and skills in simple familiar situations with support
- ability to identify and interpret information from simple sources with guidance
- ability to communicate simple ideas briefly

達到本等級考生的典型表現如下：

- 對課程內容有初步的認識和理解，在協助下，能把概念和技巧應用到簡單熟悉的情境
- 在指導下，能辨識和詮釋簡單的資料
- 能粗略地傳達簡單意念



# 4 The grading procedures

## 評級程序

The main purpose of grading is to determine the minimum score needed for a candidate to attain a given level. This minimum score is known as the cut score (Figure 1).

評級主要的目的是釐定考生須至少獲得多少分才達到某個等級。每個等級的最低分數稱為臨界分數(圖一)。

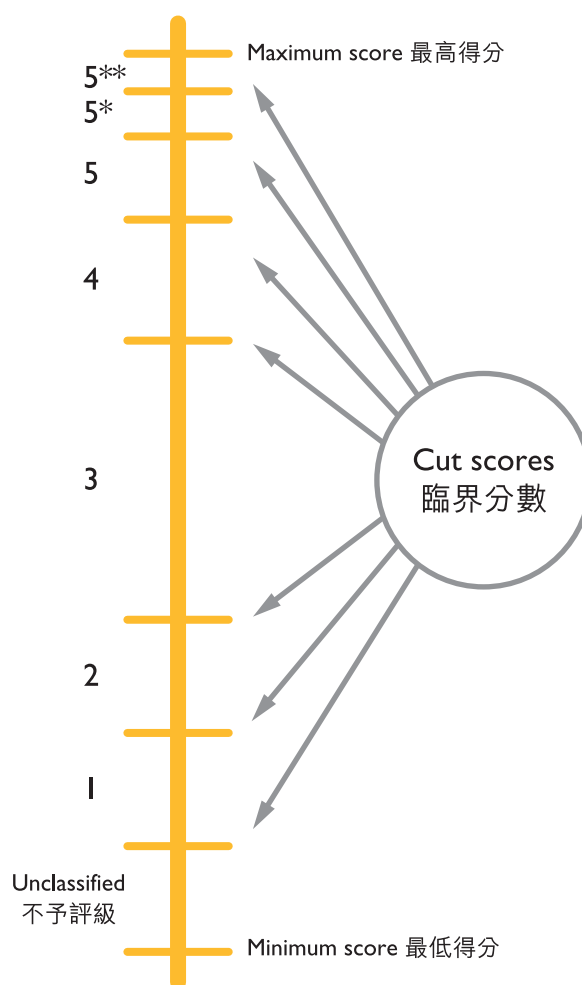


Figure 1 Cut scores  
圖一 臨界分數

The HKDSE grading procedures include a series of tasks that begins before the actual marking of scripts. For any given subject, a panel of expert judges consisting of the subject manager(s), the chief examiners and selected assistant examiners or markers from the individual components is responsible for the implementation of the following grading procedures.

### Sample script selection

After the examination, some samples that panel members feel illustrate performance particularly well in relation to the level descriptors will be selected.

### Standardisation

After script selection, the panel will discuss the marks to be awarded to discrete points in the sample scripts. They will also make observations on any perceived changes in standard from the previous year's work. These marked scripts will be used as standardisation scripts for marking.

### Post-marking exercise

After completion of marking and moderation of School-based Assessment (SBA) scores (if appropriate), the panel will discuss the standard requirement of each level based on exemplars and/or observations from the live examination and SBA sample review. The objective of the discussion is to make provisional grading recommendations (including preliminary cut score ranges) on each examination paper or component through expert judgment based on samples of performance. Details regarding standards setting and maintenance are explained in sections 5 and 6.

香港中學文憑考試的評級程序涉及的一系列工作會在正式閱卷前開始。每一科目的評級工作將由一專家小組負責執行，小組成員包括科目經理、試卷主席、經甄選的助理試卷主席或相關的閱卷員。具體程序如下：

### 選取樣本答卷

考試完成後，小組成員將選取最能符合等級描述的答卷作為樣本答卷。

### 標準化

選取樣本答卷後，小組將討論樣本答卷的評分和具體得分點。他們也會比較當年與往年的水平，並留意兩者的差異。評閱的樣本答卷會用作閱卷時參考的標準答卷。

### 閱卷後程序

完成閱卷和校本評核分數調整（如適用）後，小組會檢視公開考試和校本評核樣本的示例，並綜合觀察所得，討論各等級水平的要求。有關討論讓專家小組能針對樣本的表現作出判斷，為各試卷或分部作初步的評級建議（包括初步臨界分數範圍）。有關設定水平和維持水平的詳情，可參考第5章和第6章。

## Panel of judges grading meeting

Management team members from the Assessment Development Division and the Assessment Technology and Research Division of the HKEAA will meet with the panel judges and agree on cut scores for each paper/component and for the subject. If there is any discrepancy between their recommendations and the statistical indicators, the panel's expert recommendations will be given appropriate weighting in this exercise. The statistical indicators should, therefore, be used to check and verify the reasonableness of the cut scores proposed by the panel of judges for each subject. Any significant discrepancies between the panel's recommendations and the statistical indicators should be fully discussed and if they cannot be resolved, the rationale behind the panel's recommendations should be presented. During this meeting, the panel of judges can investigate the impact of amending the cut scores for each examination paper on subject grade distributions. By the end of the meeting for each subject, the panel of judges will decide on their recommendations for the cut scores for that subject.

## Internal meeting

A senior management team led by the Secretary General of the Authority will meet to review the recommendations made by individual subject panels. In cases where significant discrepancies remain between the recommendations of the judge panel and the statistical indicators after the panel of judges grading meetings, this internal meeting will make overriding recommendations to the Public Examinations Board on the appropriate alignment between statistical indicators and the subject experts' recommendations for each subject.

## Public Examinations Board meeting

The recommendations of the internal meeting will be reviewed at the Public Examinations Board meeting. The final cut scores of all HKDSE subjects must be endorsed at this meeting.

## 專家小組評級會議

考評局評核發展部與評核科技及研究部管理小組的成員，會與專家小組商議，就各個科目及其分部的臨界分數達成共識。當統計分析結果建議的臨界分數與專家建議的臨界分數出現明顯差異時，將會適當考慮專家的意見，而統計結果則作為參考，以驗證臨界分數的合理性。如未能達成共識，專家小組將全面討論不同的意見，並記錄其理據。專家小組可在會議中探討修訂各卷的臨界分數對整體等級分布的影響。最後，專家小組會建議有關科目的臨界分數。

## 內部會議

以考評局秘書長為首的高層管理小組，會在內部會議中審視專家小組的評級建議。如統計數據與專家小組建議出現分歧，並無法於專家小組評級會議取得共識，內部會議將適當地平衡統計數據與專家小組的建議，對有關科目的臨界分數作出凌駕性建議，供公開考試委員會審定。

## 公開考試委員會會議

公開考試委員會將審定內部會議的建議，並通過所有科目的最終臨界分數。

# 5 Setting the standards 設定水平

There are various methods that can be used to set standards. The HKEAA adopts a strategy that combines the benefits of both expert judgment and statistical methods. Experts make use of statistical recommendations and other reference data to make the most robust decisions possible in cut score determination. In principle, the standards of all HKDSE subjects in the 2012 examination will be set with reference to the level descriptors and actual performance of candidates in the examination. To support experts' decision making, different statistical techniques will be applied to different subjects to produce recommended cut scores. The methodologies are as follows:

## Core subjects

In setting standards for the four HKDSE core subjects – Chinese Language, English Language, Mathematics and Liberal Studies, reference will be made to the actual performance of candidates in the 2012 HKDSE Examination in relation to:

- the level descriptors;
- the standards of candidates in the 2011 Hong Kong Advanced Level (HKAL) Examination;
- the standards of candidates in the 2010 Hong Kong Certificate of Education (HKCE) Examination.

In order to provide some background information about standards in the HKDSE and previous HKCE/HKAL examinations, secure monitoring tests for the core subjects were administered to samples of HKCE school candidates in 2010 and HKAL school candidates in 2011. The same monitoring tests will also be administered to a selected sample of candidates sitting the 2012 HKDSE Examination. In the first administration, the statistical data from the monitoring tests will help provide some linkage of the standards in the HKDSE and the previous HKCE and HKAL examinations.

The panel of judges will refer to the following information in making their decisions on cut scores:

- marked live scripts, selected according to total marks;
- inter-paper correlations, the mean and standard deviation of the current year's papers;
- paper mark cumulative distributions;
- feedback from markers on the level of difficulty of each particular examination paper;
- performance samples from the HKDSE SRR Information Packages;
- HKAL and HKCE papers and library scripts from 2011 and 2010 examinations respectively.

設定水平的方法有很多，而考評局採用一套結合專家判斷和統計技術的方法。當中專家會參考統計模型的建議及其他參考資料，釐定最合適的臨界分數。原則上，2012年香港中學文憑考試各科的水平會按照等級描述和考生實際表現設定。不同科目會利用不同的統計方法計算建議臨界分數，以供專家參考。有關統計方法如下：

## 核心科目

香港中學文憑考試四個核心科目，即中國語文、英國語文、數學及通識教育，這些科目的水平設定將參考2012年考生實際表現與以下各項的關係：

- 等級描述；
- 2011年香港高級程度會考(高考)考生水平；
- 2010年香港中學會考(會考)考生水平。

為提供香港中學文憑考試和以往高考與會考三者的水平的背景資料，考評局分別替部分參與2010年會考和2011年高考的學校，進行保密研究測驗。同樣的研究測驗亦將在參與2012年香港中學文憑考試的部分學校中進行。這些從研究測驗取得的統計數據，可作為連結首屆香港中學文憑考試和以往高考與會考水平的參考資料。

專家小組在釐定臨界分數時，也會參考以下資料：

- 按總分選取的已評閱答卷；
- 該年的試卷間相關係數、平均分及標準差；
- 試卷總分的累積分布；
- 閱卷員就個別試卷難易程度的回饋意見；
- 水平參照成績匯報資料套內不同能力水平的樣本；
- 2010年會考和2011年高考的試卷和存檔的參考答卷。

## Elective subjects

A group ability index (GAI) will be used as a reference for grading elective subjects. The GAI is a set of percentages generated statistically for obtaining a suggested set of cut scores. For a group of candidates taking a specific elective subject, a GAI can be calculated for each level based on the number of candidates in this group achieving that level in the four core subjects. The method of GAI calculation is detailed in Appendix 1.

In the first administration of the HKDSE Examination in 2012, the GAI-based cut scores are only for reference. The panel of judges for each subject will play a very important and independent role in setting a cut score for each level based on the actual performance of candidates.

## Applied Learning subjects

Standards of 'Attained' are initially determined by course providers and moderated by panels of judges with reference to the performance descriptors of individual subjects. Candidates awarded 'Attained with Distinction' are deemed to have performed at a level comparable to Level 3 or above for Category A subjects. Therefore, the award percentage of 'Attained with Distinction' will be determined based on the GAI for Level 3 or above.

## Special arrangements for individual subjects

Several HKDSE subjects consist of elective modules/parts. In order to set common standards for candidates choosing different modules/parts of the same subject, it is necessary to equate marks in different modules/parts. A method of equipercentile mark conversion will be used for this purpose. The principle and application of this method are explained in Appendix 2.

## 選修科目

考評局會利用「組別能力指數」作為選修科目評級的參考。「組別能力指數」是利用統計方法計算出的一組百分比，以得到一組建議臨界分數。對於應考某選修科目的全體考生來說，他們在核心科目取得某等級的人數，會用來計算出該選修科目於相應等級的「組別能力指數」。有關「組別能力指數」的計算方法詳列於附錄1。

在2012年首屆香港中學文憑考試，「組別能力指數」建議的臨界分數只會作參考之用，各科的專家小組可按考生的實際表現，獨立為各級水平釐定合適的臨界分數。

## 應用學習科目

課程提供者先設定「達標」的水平，再由專家小組按個別科目的表現描述作調整。由於「達標並表現優異」的水平等同甲類科目的第3級或以上，因此獲得「達標並表現優異」水平的考生比例，將參考「組別能力指數」第3級或以上的百分比而釐定。

## 個別科目的特殊安排

香港中學文憑考試部分科目設有選修部分。為劃一不同選修部分考生的水平，考評局採用百分位等值法為不同選修部分作分數等值。該方法的原理和應用載於附錄2。

## Determination of Level 5\*\* and Level 5\*

The cut scores for Level 5\*\* and Level 5\* will be set with reference to the percentage in mark distribution so that Level 5\*\* will be awarded to the highest-achieving 10% (approximately) of Level 5 candidates and Level 5\* will be awarded to the next highest-achieving 30% (approximately) of Level 5 candidates.

For Chinese Language and English Language, the overall subject cut score for Level 5\*\* and Level 5\* will be set first. The corresponding cut scores for each component will then be determined by exploration and negotiation across the components in such a way that they represent a similar profile of division within Level 5 for each component. However, the cut scores established for each component will still aggregate to achieve the correct corresponding subject cut scores for Level 5\*\* and Level 5\*.

For Combined Science, the cut scores of the individual components of Biology, Chemistry, and Physics will be determined with reference to the cut scores of the corresponding subjects using the equipercentile method. The cut score for Level 5\*\* and Level 5\* for each half subject will be determined by applying the above-mentioned percentages to Level 5 candidates. The half subject cut scores will then be aggregated to obtain the subject cut scores for each subject pairing of Combined Science. In other words, the respective percentage of Level 5 candidates attaining subject levels 5\*\* and 5\* in Combined Science may not necessarily be 10% and 30% as in other subjects.

## 5\*\*級與5\*級的釐定

5\*\*級和5\*級的臨界分數是按第5級內考生的分數分布而訂定。其中，約10%成績最優異的第5級考生，可獲5\*\*級，約30%成績次佳的第5級考生，可獲5\*級。

至於中國語文和英國語文科，專家小組將先釐定整個科目5\*\*級和5\*級的臨界分數，然後推算和釐定個別分部的臨界分數，使各分部的5\*\*級與5\*級的分布較為一致；然而，各分部的臨界分數相加後，仍須相等於整個科目相應等級的臨界分數。

就組合科學而言，生物、化學及物理等分部的臨界分數將採用百分位等值法，參考相應的三個科目的臨界分數而釐定。組合科學的分部科目的5\*\*級和5\*級臨界分數，會按其第5級的分數分布，根據上述百分比釐定。就每兩個分部科目組合而言，把有關的兩個分部科目的臨界分數加起來便得出整科的臨界分數。故此，在組合科學科，整體科目獲第5級的考生，能取得5\*\*級和5\*級的人數比例並不一定如其他科目那樣定為10%和30%。

# 6 Maintaining the standards 維持水平

Once set, the standards need to be maintained in subsequent years starting from 2013 so that results are comparable across years.

Similar to the standards setting procedures, panel judges of different subjects will refer to various kinds of information for the determination of cut scores, including:

- marked live scripts, selected according to total marks;
- inter-paper correlations, the mean and standard deviation of the current year's papers in comparison to previous years;
- item statistics for previous years for reference;
- paper mark cumulative distributions;
- feedback from markers on the level of difficulty of each particular examination paper;
- HKDSE library scripts.

The panel judges will also refer to results generated by statistical models. The statistically-generated suggested cut scores for the core subjects and those for the other subjects will be computed using different methods.

## Core subjects

For the four core subjects, secure monitoring tests are conducted annually in a representative group of carefully selected schools. Using both monitoring test data and live examination data as input, a latent trait model is used to calibrate all questions in different examinations and generate suggested cut scores. The details of a latent trait model are documented in Appendix 3.

As shown in Figure 2, the statistically-generated suggested cut scores for the 2013 examinations will be obtained by mapping the cut scores established for the 2012 examinations to the 2013 live examination data with the standard maintained by means of test equating techniques using the results from the monitoring tests. In this way, the statistical equating of successive examinations in the core subjects will support the panel judges in making their final decisions on the recommended cut scores, using the grading procedures described in section 4.

水平標準一經制訂，考評局須從2013年開始維持往後考試的水平，以確保年與年之間的成績有可比性。

與水平設定的程序相若，不同科目的專家小組會參考各種資料以釐定臨界分數。這些資料包括：

- 按總分選取的已評閱答卷；
- 該年的試卷間相關係數、平均分及標準差與往年的比較；
- 往年各分題的統計數據；
- 試卷總分的累積分布；
- 閱卷員就個別試卷難易程度的回饋意見；
- 香港中學文憑考試存檔的參考答卷。

專家小組亦會參考統計模型建議的臨界分數。核心科目和其他科目會採用不同統計模型來計算建議臨界分數。

## 核心科目

考評局每年均會小心選取一些學校，為四個核心科目進行保密研究測驗。根據研究測驗與該年公開考試的數據，運用「潛在特性模型」校定不同考試的題目的難度，並以此計算建議臨界分數。有關「潛在特性模型」的詳情，可參考附錄3。

如圖二所示，運用研究測驗結果，通過等值方法，將2013年度和2012年度的考試等值。而根據2012年度考試設定的臨界分數，找到2013年度考試的等值對應分數，就是統計建議的臨界分數。這樣，專家小組可參考年與年之間核心科目考試成績的等值結果，再根據第4章的程序，釐定最終的建議臨界分數。

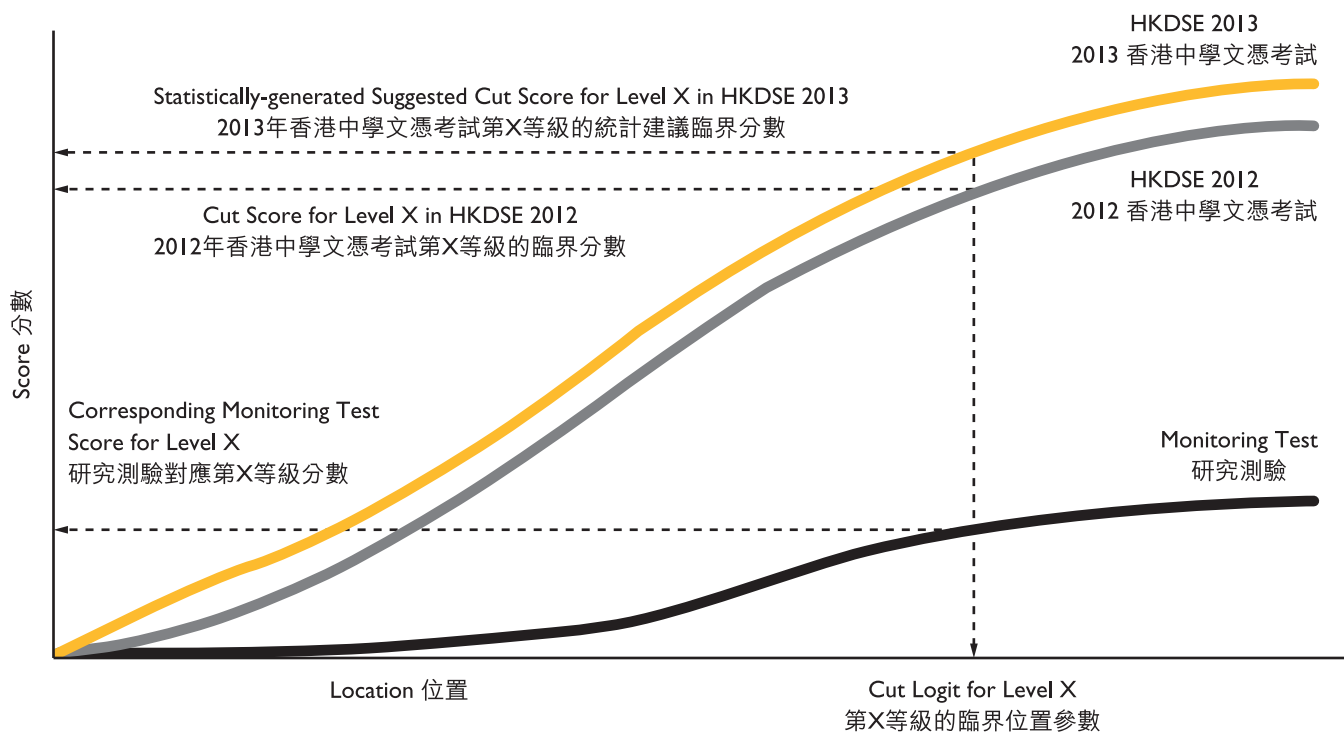


Figure 2 Using monitoring tests in maintaining standards and producing suggested cut scores for the core subjects  
圖二 利用研究測驗來維持核心科目水平及產生建議臨界分數



## Elective subjects

Recommended cut scores for the elective subjects will be suggested using the GAI. Similar to the grading of the core subjects, panel judges for the elective subjects may then consider these suggested cut scores and use them to support their final recommendations on standards-maintenance.

## Applied Learning subjects

The standards of 'Attained' are maintained by making reference to the performance descriptors of individual subjects and the actual performance of candidates in the previous years. For 'Attained with Distinction', the GAI will be used to make suggestions for the proportion of candidates to be awarded this level by referring to the same group of candidates who have attained Level 3 or above in the core subjects.

## Other language subjects

The standards of these subjects will be maintained by CIE since their AS level papers are used for the examinations of these subjects.

The above standards-maintenance processes will assure the public, local and overseas tertiary institutions, employers, and other users that the standards are held constant and there is no 'grade inflation' over time.

## 選修科目

選修科目的建議臨界分數是依據「組別能力指數」計算的。與核心科目的評級方法相若，選修科目的專家小組可參考以「組別能力指數」計算所得的建議臨界分數，並藉此作最終的建議，以維持各等級的水平。

## 應用學習科目

為了維持「達標」水平，考評局會參照個別科目的表現描述和往年考生的真實表現。考生達到「達標並表現優異」等級的比例，將參照該組別考生於核心科目中達第3級或以上的「組別能力指數」而釐定。

## 其他語言科目

由於這些科目的考試採用英國劍橋大學國際考試組高補程度的試卷，故有關科目的水平亦由該機構維持。

以上維持水平的方法可以向公眾、香港及海外的大專院校、僱主及其他使用者確保各等級的水平標準是固定的，日後亦不會出現「等級膨脹」的情況。

# 7 Reporting the results 匯報成績

In order to provide more detailed information, the HKDSE certificate will list both subject level results as well as component level results (if any). SBA is an integral part of Category A subjects, so SBA results are not reported separately. Moderated SBA results (if any) are combined with public examination results to form an overall component or subject level as appropriate. A draft example of how results will be set out on the certificate is given in Figure 3.

為提供更詳細的資料，香港中學文憑成績證書將列出科目及分部(如有)的成績等級。校本評核是甲類科目不可分割的部分，因此校本評核的成績將不會獨立匯報。經過調整的校本評核分數(如有)將與公開考試成績合併，並在分部或科目等級中一併匯報。考生成績證書的初稿如圖三所示。

**香港考試及評核局**  
HONG KONG EXAMINATIONS AND ASSESSMENT AUTHORITY  
**香港中學文憑考試**  
HONG KONG DIPLOMA OF SECONDARY EDUCATION EXAMINATION

茲證明  
This is to certify that

**陳大文**  
CHAN TAI MAN

參加 2012 年香港中學文憑考試，成績如下：  
sat the 2012 Hong Kong Diploma of Secondary Education Examination and achieved the following results:

科目 Subject	科目等級 Subject Level / Grade	分部等級 Component Level
<b>甲類 CATEGORY A</b>		
中國語文 · 閱讀 · 寫作 · 聆聽 · 說話 · 綜合能力	CHINESE LANGUAGE · Reading · Writing · Listening · Speaking · Integrated Skills	5(Five)  5(Five) 5(Five) 5(Five) 5(Five)
英國語文 · 閱讀 · 寫作 · 聆聽與綜合能力 · 說話	ENGLISH LANGUAGE · Reading · Writing · Listening and Integrated Skills · Speaking	5**(Five**)  5**(Five**) 5**(Five**) 5**(Five**)
數學 必修部分	MATHEMATICS Compulsory Part	5*(Five*)
通識教育	LIBERAL STUDIES	5(Five)
組合科學 · 生物 · 化學	COMBINED SCIENCE · Biology · Chemistry	4(Four)  4(Four) 4(Four)
<b>乙類 CATEGORY B</b>		
運動科學及體適能	EXERCISE SCIENCE AND HEALTH FITNESS	達標並表現優異 Attained with Distinction
<b>丙類 CATEGORY C</b>		
法語	FRENCH	A(a)

---

香港身分證號碼/旅遊證件號碼  
Hong Kong Identity Card No./ : Z123456(7)  
Travel Document No.

考生編號  
Candidate No. : 121234568

註釋見背頁  
Explanatory Notes are printed overleaf

證書編號  
Certificate No. : **0123456**

秘書長 Secretary General

Figure 3 A draft of the HKDSE certificate  
圖三 香港中學文憑證書初稿

# Appendix I : Group Ability Index (GAI)

## 附錄 I : 組別能力指數

Group Ability Index (GAI) can be regarded as a set of 'suggested' percentages that will be used as a reference for the grading of elective subjects and Applied Learning subjects.

The formula for calculating the Group Ability Index  $P$  of Subject  $X$  for a certain level or above (e.g. Level 3 or above) is as follows:

$$P = \frac{1}{b_C + b_E + b_M + b_L} \left( b_C \frac{n_C}{N_C} + b_E \frac{n_E}{N_E} + b_M \frac{n_M}{N_M} + b_L \frac{n_L}{N_L} \right) \times 100\%$$

where  $N_C$ ,  $N_E$ ,  $N_M$  and  $N_L$  are the numbers of Subject  $X$  candidates who sat the four core subjects – Chinese Language, English Language, Mathematics (Compulsory Part) and Liberal Studies, respectively;

$n_C$ ,  $n_E$ ,  $n_M$  and  $n_L$  are the numbers of Subject  $X$  candidates who earned that level or above (i.e. Level 3 or above) in the four core subjects of Chinese Language, English Language, Mathematics (Compulsory Part) and Liberal Studies, respectively;

$b_C$ ,  $b_E$ ,  $b_M$  and  $b_L$  are coefficients obtained by regressing the standardised scores of Subject  $X$  on the standardised scores of the four core subjects.

After calculating the value of  $P$ , the suggested cut score for that level or above (i.e. Level 3 or above) is the  $1-P$  percentile of the Subject  $X$  scores.

「組別能力指數」基本上可視作一組百分比，用作選修科目和應用學習科目評級的參考。

計算某科目  $X$  某個等級或以上(例如第 3 級或以上)的組別能力指數  $P$  的公式如下：

$$P = \frac{1}{b_C + b_E + b_M + b_L} \left( b_C \frac{n_C}{N_C} + b_E \frac{n_E}{N_E} + b_M \frac{n_M}{N_M} + b_L \frac{n_L}{N_L} \right) \times 100\%$$

其中  $N_C$ 、 $N_E$ 、 $N_M$  及  $N_L$  分別代表科目  $X$  的考生應考四個核心科目：中國語文、英國語文、數學(必修部分)及通識教育科的總人數；

$n_C$ 、 $n_E$ 、 $n_M$  及  $n_L$  分別代表應考某個科目  $X$  的考生在中國語文、英國語文、數學(必修部分)及通識教育科四個核心科目中獲得某個等級或以上(例如第 3 級或以上)的人數；

$b_C$ 、 $b_E$ 、 $b_M$  及  $b_L$  是把科目  $X$  的標準分與四個核心科目的標準分作回歸分析，得出的回歸係數。

計算出  $P$  的數值後，某個等級或以上(例如第 3 級或以上)的建議臨界分數為科目  $X$  分數的  $1-P$  百分位點。

## Appendix 2 : Equipercentile method

### 附錄 2 : 百分位等值法

In some examination papers of HKDSE subjects, there is one compulsory part and two or more optional parts. Equating is needed so that the performance of candidates choosing different optional parts can be reflected on the same scale.

The idea of equating is to convert the marks of one optional part into another optional part by using the compulsory part as a mediator, or to convert marks of all optional parts into the marks of the compulsory part. This can be done in three steps:

- Ranking the candidates according to their performance in each part;
- Dividing the candidates of each part into equal-sized groups (for example 20 groups of candidates, each made up of 5% of the total candidature);
- Converting the marks of one part to the marks of another part with reference to the mark distribution in different groups; if necessary, making adjustments in cases with the same percentile.

For example, in the subject of English Language, there are four papers in the public examination, namely, Reading (Paper 1), Writing (Paper 2), Listening and Integrated Skills (Paper 3) and Speaking (Paper 4). All parts in Writing and Speaking are compulsory. For Reading and Listening and Integrated Skills, however, only Part A is compulsory. All candidates must do Part A and choose either Part B1 (easier section) or Part B2 (more difficult section).

To assess the performance of candidates choosing different sections on the same scale, the marks of the easier section will be converted to a scale comparable to the one used for the more difficult section. This will be done using the marks of the compulsory section as the mediator. The resulting marks will be known as the equivalent difficult section marks. A candidate's total paper marks will then be calculated by adding the marks of the compulsory section to either the marks of the more difficult section or the equivalent difficult section marks. Grading can then be done based on total paper marks for all candidates.

部分香港中學文憑考試的試卷包括一個必答部分，和兩個或多個選答部分。因此，考評局需要進行分數等值，使選答不同部分考生的表現，能以同一尺度作比較。

分數等值的概念是把某選答部分的分數，以必答部分的分數作中介，轉換為另一選答部分的分數。又或是把不同選答部分的分數都轉換為必答部分的分數。分數等值涉及三個步驟：

- 按考生表現把每個部分的分數排序；
- 把每一部分的考生分為多個等份（例如分為 20 等份，每等份佔全部考生 5%）；
- 按各等份的分數分布，把某一部分的分數轉換為另一部分的分數；如有需要，調整相同百分位點的分數。

以英國語文科為例，公開考試部分包括四份試卷：分別是閱讀（卷一）、寫作（卷二）、聆聽與綜合能力（卷三）及說話（卷四）。寫作和說話卷中所有部分全屬必答。而閱讀和聆聽與綜合能力卷，則只有 A 部分為必答题，考生須作答 A 部分的必答题和選答 B1 部分（較易部分）或 B2 部分（較難部分）的題目。

為了以劃一的尺度評估選答不同部分考生的表現，考評局將參考必答部分的分數分布，把較易部分的分數轉換到一個與較難部分可比的尺度。轉換後的結果稱為較難部分的等值分數。把每一考生必答部分的分數和較難部分的分數（或較難部分的等值分數）相加便得出試卷總分，繼而根據試卷總分進行評級。

## Appendix 3: Latent trait model

### 附錄 3：潛在特性模型

A latent trait model developed from modern item response theories is a statistical tool for the calibration of examination items and the estimation of candidates' ability levels based on their responses to items in a test.

The model that will be used to calibrate the HKDSE examinations is based on the Rasch model. The simple Rasch model for dichotomous (i.e., either right or wrong) responses is:

$$p_{ni} = \Pr\{X_{ni} = 1\} = \frac{\exp(\beta_n - \delta_i)}{1 + \exp(\beta_n - \delta_i)} \quad (1)$$

However, most examinations contain questions that are not simply right or wrong, and for which more than one mark may be awarded. The model for such polytomous responses is:

$$\Pr\{X_{ni} = x_{ni}\} = \frac{\exp\{x_{ni}(\beta_n - \delta_i) - \sum_{k=0}^{x_{ni}} \tau_{ik}\}}{\sum_{l=0}^m \exp\{l(\beta_n - \delta_i) - \sum_{k=0}^l \tau_{ik}\}} \quad (2)$$

where  $\beta_n$  is the ability of the person  $n$ ,  $\delta_i$  is the difficulty of the item  $i$ ,  $\{\tau_{ik}\}$  are the centralised thresholds and  $\sum_{k=0}^0 \tau_{ik} = 0$ .  $\beta_n$  and  $\delta_i$  are generally called the locations (logits).

This is the model that will be used to calibrate the items of live examinations together with the data of the monitoring test of a subject. To equate the examination scores across different years, the following formula will be adopted to calculate the total score of a candidate on any subset of items:

$$SubS = \sum_{i \in I} K_i E(X_i) \quad (3)$$

「潛在特性模型」(latent trait model)是建基於現代題目反應理論的統計工具，能通過分析考生對題目的反應(考試表現)校定考試題目及估計學生的能力水平。

校定香港中學文憑考試題目的「潛在特性模型」是根據 Rasch 模型發展出來的。簡單的二值(即只有「對」或「錯」兩個答案)反應 Rasch 模型如下：

$$p_{ni} = \Pr\{X_{ni} = 1\} = \frac{\exp(\beta_n - \delta_i)}{1 + \exp(\beta_n - \delta_i)} \quad (1)$$

但是，大部分考試題目都不是簡單的只有「對」或「錯」兩個答案的問題，每題答對的得分亦可能多於一分。這時，需要以下的多級評分模型：

$$\Pr\{X_{ni} = x_{ni}\} = \frac{\exp\{x_{ni}(\beta_n - \delta_i) - \sum_{k=0}^{x_{ni}} \tau_{ik}\}}{\sum_{l=0}^m \exp\{l(\beta_n - \delta_i) - \sum_{k=0}^l \tau_{ik}\}} \quad (2)$$

其中  $\beta_n$  是考生  $n$  的能力， $\delta_i$  是題目  $i$  的難度， $\{\tau_{ik}\}$  是已經中心化的閾值 (centralised thresholds)，而  $\sum_{k=0}^0 \tau_{ik} = 0$ 。 $\beta_n$  與  $\delta_i$  則統稱為位置參數 (logits)。

這是考評局將用作校定某科目公開考試題目與研究測驗題目的模型。為把不同年份的考試分數等值，考生在某科目任何題目子集的總分將按以下公式計算：

$$SubS = \sum_{i \in I} K_i E(X_i) \quad (3)$$

where  $K_i$  is the pre-specified multiplying constant of item  $i$ ,  $I$  is the subset of items in the subject, and  $E(X_i)$  is the expected value for the candidate  $n$  with ability  $\beta$  on item  $i$  with the maximum score of  $m_i$

$$E(X_i) = \sum_{x'=1}^{m_i} x' P(X_i = x' | \beta) \quad (4)$$

Calibration of all items on different examination papers and the secure monitoring tests enables estimates of the ability of each candidate to be generated on the same scale.

After calibrating item difficulties, it is possible to compile the expected score of a whole subject (which may consist of different components) for a given candidate's ability by considering the probability of that candidate obtaining a certain mark and different weightings for the different components. These weightings are calculated based on the structure of the examination papers and the performance of the candidates in the examination.

This model supports the grading work of panel judges by generating suggested cut scores for each level of the four core subjects for their reference.

Starting from 2013, in order to provide basic reference data for equating successive examinations and thus enable standards to be maintained from one year to the next, the HKEAA will make use of the above latent trait model to create an overall calibration of responses to all questions on each examination paper (over two particular years) and all questions on the secure monitoring tests.

其中  $K_i$  是題目  $i$  的設定權重， $I$  是該科目的題目子集，而  $E(X_i)$  是某一能力水平為  $\beta$  的考生  $n$  在某一滿分為  $m_i$  分的題目  $i$  的預期分數

$$E(X_i) = \sum_{x'=1}^{m_i} x' P(X_i = x' | \beta) \quad (4)$$

通過同時校定各份試卷和保密研究測驗的所有題目，便可以用相同尺度估計每個考生的能力。

當校定試題的難度後，便可為某一考生能力值計算相對應的整個科目（亦可能包括不同分部）的分數期望值。在計算過程中，會顧及取得某個分數的概率，以及科目考試可能包括多個獨立分部，每一分部有不同的比重。該比重是根據試卷結構及考生表現等資料計算出來的。

本模型產生四個核心科目各等級的建議臨界分數，供專家小組評級時參考。

從 2013 年開始，為提供能將連續多個年度考試等值的基本參考資料，考評局將採用以上的「潛在特性模型」校定（某兩年）公開考試每份試卷中所有題目與保密研究測驗中所有題目的參數，以作為維持每年等級水平的參考。

## Reference materials

### 參考資料

Details of the level descriptors for different subjects can be found in the HKDSE Standards-referenced Reporting Information Packages for individual subjects. These are available for sale at the HKEAA sub-offices in Wan Chai and San Po Kong, as well as the online bookstore.

<http://www.hkeaa.edu.hk/en/Resources/publications/>

個別科目的《香港中學文憑水平參照成績匯報資料套》載有該科等級描述的詳情。這些資料套於考評局灣仔與新蒲崗辦事處及網上書店發售。

<http://www.hkeaa.edu.hk/tc/Resources/publications/>

Further information about grading procedures and standards-referenced reporting can be obtained from the HKEAA website at

如須進一步瞭解評級程序與水平參照成績匯報的詳情，可瀏覽考評局網頁

**[www.hkeaa.edu.hk](http://www.hkeaa.edu.hk)**

Enquiries can be directed to:

如有查詢，請聯絡：

**Hong Kong Examinations and Assessment Authority**

**香港考試及評核局**

Address 地址：13/F, Southorn Centre, 130 Hennessy Road, Wan Chai, Hong Kong

香港灣仔軒尼詩道一百三十號修頓中心十三樓

Tel 電話：(852) 3628 8860

Fax 傳真：(852) 3628 8928

Email 電郵：dse@hkeaa.edu.hk



