

香港中學文憑考試

訂定甲類科目重閱答卷後提升等級所需特定分數的方法

背景

香港中學文憑考試（文憑試）的積分覆核及重閱答卷政策旨在保障全部考生，不論其有否申請積分覆核及重閱答卷，均獲得可靠與公平的等級。該政策包括提升等級的準則。香港考試及評核局（考評局）會定期檢討該政策，任何修訂建議，均會交予公開考試委員會考慮和通過。

積分覆核及重閱答卷後提升等級的準則

1. 積分覆核時如發現技術錯誤，修正後的最終分數只要達到更高等級的臨界分數，科目／分部等級便會獲得提升。
2. 重閱答卷後，最終分數須達到在更高等級的臨界分數之上的特定分數，科目／分部等級方會獲得提升。

設定提升等級的特定分數要求的理據

考評局致力完善閱卷與評級的各個環節，以確保所頒授等級的準確度。有關閱卷與分數處理的嚴格程序，可參考網頁：www.hkeaa.edu.hk/tc/hkdse/assessment/marking/

甲類科目的等級是由專家小組在每年考試後按水平參照評級程序所訂定，並由公開考試委員會審議和通過。有關文憑試的評級程序與水平參照成績匯報的詳情，可參閱考評局網站的資料：www.hkeaa.edu.hk/tc/HKDSE/assessment/the_reporting_system/SRR/

在成績發放後，部分考生會基於種種原因申請積分覆核及重閱答卷。考評局會首先覆核申請重閱答卷考生的答卷所得積分是否有技術錯誤，然後交由覆核閱卷員重閱答卷，並盡量確保重閱答卷的閱卷員，並非原先的閱卷員。

然而，閱卷工作涉及專業判斷，特別是評閱開放式題目。由於重閱答卷的閱卷員群組與原來的群組不盡相同；加上重閱答卷的時間和原來閱卷亦相距數月，因此重閱答卷的得分可能有別於原來得分。有關評核情境的變動，會導致水平相若的答卷得分出現差異。由於有這種自然和合理的分數變動，假若某考生經重閱答卷後的最終分數僅達更高等級的臨界分數，這並不足以證明原來等級未能準確反映考生的成績，該考生有可能實際上未達到更高等級的水平。因此，在重閱答卷的程序中，必須有強而有力的顯證來支持原來的等級需要改變，以保障全體考生所得等級的整體可靠度。

因此，重閱答卷後，最終分數須達到在更高等級的臨界分數之上的特定分數，科目／分部的等級方會獲得提升。這個重閱答卷後分數變動的容許界限，確保考生無論有否申請重閱答卷，都會公平地獲得應有的等級。

訂定提升等級的特定分數的方法

基於上述考慮和教育測量理論，考評局為所有甲類科目各個等級制定提升等級政策，重閱答卷的最終科目分數必須高出更高等級的臨界分數若干分(即特定分數)，考生方能獲得等級提升。

特定分數以「標準分」為單位，而一標準分的定義為每個相關等級的分數範圍的二十分之一。每個科目的等級特定分數都設定在相同數量的標準分，但其科目分數則受以下因素影響：

1. 更高等級所涵蓋的分數範圍

個別科目每個等級的分數範圍取決於相鄰等級的臨界分數。臨界分數由專家小組依據水平參照評級程序訂定，並經公開考試委員會通過。

2. 閱卷涉及的科目專業判斷程度

閱卷時涉及較多整體專業判斷的科目，包括開放式題目的科目，其等級提升所要求的特定分數會較高，即要求較多的標準分；而評分時涉及較多客觀分析或科目分數範圍很小的科目，則設較低的特定分數，即要求較少的標準分。

3. 科目的總科目分數的測量標準誤 (*standard error of measurement, SEM*)

測量標準誤是測考中的一個重要元素，因為任何測量都無可避免涉及測量的誤差。它反映了同一考生在不同時間參與類似考試，或答卷在不同評核情境下由另一名閱卷員評分時可能出現的分數差異。考評局會計算個別科目的科目分數測量標準誤，並作定期監察。測量標準誤的計算詳見於附件。

文憑試各個甲類科目所有等級的特定分數，由考評局參考以上三個因素而訂定，這個重閱答卷後分數變動的容許界限是公平和合理的。特別需要注意的是考生在重閱答卷後，即使最終分數低於原來等級的臨界分數，按目前的積分覆核及重閱答卷政策，其成績也不會被降級。

考評局會定期檢討和監察特定分數是否適當。如有需要，特別是當考試形式或評核架構有重大變更時，相應的修訂建議會交予公開考試委員會考慮和通過。

附件：測量標準誤（standard error of measurement）的計算

測量標準誤反映了同一考生在不同時間參與類似考試，或答卷由另一名閱卷員評分時可能出現的分數差異。它是一個估計考生的真分數(考生真正能力／知識的一個理論數值)所在範圍的方法。

任何測量都無可避免地涉及測量誤差。簡單地說，在教育測量中，第 i 個考生在考試中獲得的分數(S_i)，即是其觀察分數。 S_i 是考生真分數(T_i)與誤差(E_i)兩者之和，即是：

$$S_i = T_i + E_i$$

一般的假設是，就總體考生而言，真分數(T)與誤差(E)在統計學上是相互獨立的變量。因此， T 、 E 與觀察分數(S)的方差 (variance, V) 會依循以下關係：

$$V(S) = V(T) + V(E),$$

而考試的信度 R_{xx} 定義為 T 的方差與 S 的方差之比：

$$R_{xx} = V(T) / V(S)$$

估計信度的數值有幾種方法，包括克隆巴赫(Cronbach's alpha)係數和試題區辨指標 (person separation index)。

在教育測量文獻中，SEM 一般是通過信度計算出來的：

$$SEM = \sqrt{V(S)(1 - R_{xx})},$$

其中 $\sqrt{V(S)}$ 是考試的標準差。