

大考中心對 108 學測試題鑑別度相關議題的說明

108.05.10

財團法人大學入學考試中心基金會大學入學考試中心（簡稱大考中心或本中心），接受大學招生委員會聯合會（簡稱招聯會）委託，辦理各項大學入學考試。108 學年度學科能力測驗（簡稱學測）已於 108 年 1 月 25 至 26 日於全國 33 考區、89 分區、3,409 個試場辦理完成，計有 138,387 位考生報考，各科選考人數與百分比如表 1：

表 1 108 學測各科選考人數與百分比

考科	選考人數	百分比
國文	138,244	99.90%
英文	138,068	99.77%
數學	136,376	98.55%
社會	124,225	89.77%
自然	111,772	80.77%

本中心於 2 月 25 日公布考生成績及各科級分人數分布、五項標準，並於 2 月 26 日寄送成績通知單。

自 91 年開始實施「考招分離」後，本中心即專責辦理大學入學考試，擔負連結大學選才與高中育才之責任，並與社會各界進行溝通，自我檢視、力求精進。惟此次學測成績公告後，特別是在個人申請第一階段篩選公告後，社會各界意見紛歧，尤其對於考試之「鑑別度」、「穩定度」、「級分夠不夠用」等議題頗多意見，其中固有積極督促之意，但也有誤解的觀點，本中心基於職責所在，認為應就其中與考試專業有關議題進行說明。

壹、鑑別度

- 一、「鑑別度」通常指考試結果是否能區分不同能力考生的程度。本中心於考後都會以考試資料進行計算與分析，並公告於本中心的網頁。分為兩部分：
 - 1.以試題為單位計算試題鑑別度，與其他統計資料一併公布於中心網頁。
 - 2.以整卷為單位計算整體鑑別度，於每次考後辦理的試題研討會中說明，並於辦理後於中心網頁公布簡報。
- 二、本中心試題鑑別度指標（D 值），是利用高分組（前 33%）考生的得分率，減去低分組（後 33%）考生的得分率。一般而言，試題鑑別度指標越高，表示鑑別度越好。

三、以各考科選擇題的平均鑑別度來看，108 學測五科由高而低，依序為數學（0.59）、英文（0.59）、自然（0.46）、國文（0.35）與社會（0.31）。進一步比較近年的平均鑑別度，國文、英文、數學三科在五年間有微幅的波動，但以 108 相較於 107 為微幅提升。108 學測數學科，無論與同年其他考科相比，或就近五年來看，平均鑑別度都不錯，此與社會議論之數學鑑別度不足的批評並不一致。

表 2 近年學測各科選擇題整體平均鑑別度

考科	108	107	106	105	104
國文	0.35	0.33	0.35	0.33	0.33
英文	0.59	0.57	0.52	0.55	0.56
數學	0.59	0.55	0.48	0.53	0.52
社會	0.31	0.37	0.35	0.33	0.36
自然	0.46	0.40	0.46	0.45	0.45

四、整卷鑑別度，可以使用顯示分散程度的得分標準差來說明。一般而言，得分標準差越大，代表考生分數的分散程度越大，即考試可以把考生分數分得比較分散，顯示整卷鑑別度越佳。得分標準差大小與分數範圍有關，因此比較時，應以分數範圍相同者進行比較，例如國文、英文、數學三科，滿分皆為 100 分。

國文、英文、數學三科中，得分標準差由高而低依序為數學（25.66）、英文（24.75）、國文（14.38），仍以數學、英文的整卷鑑別度較佳。比較近年來三科得分標準差，108 年都比 107 年微幅提升，亦與社會議論之數學鑑別度不足的批評不一致。

表 3 近年學測國文、英文與數學得分標準差

考科	得分標準差				
	108	107	106	105	104
國文	14.38	12.86	15.60	14.37	15.27
英文	24.75	24.08	22.62	22.90	23.43
數學	25.66	24.67	21.80	23.45	23.00

五、由級分人數分布圖也可看出考生成績分散程度。本次社會議論的數學科，從近年的級分人數分布圖（見圖 1）來看，108 學測的分布最低闊，不似過去幾年在左邊形成高峰（亦即，低分群較集中），鑑別度比以往相對為佳。

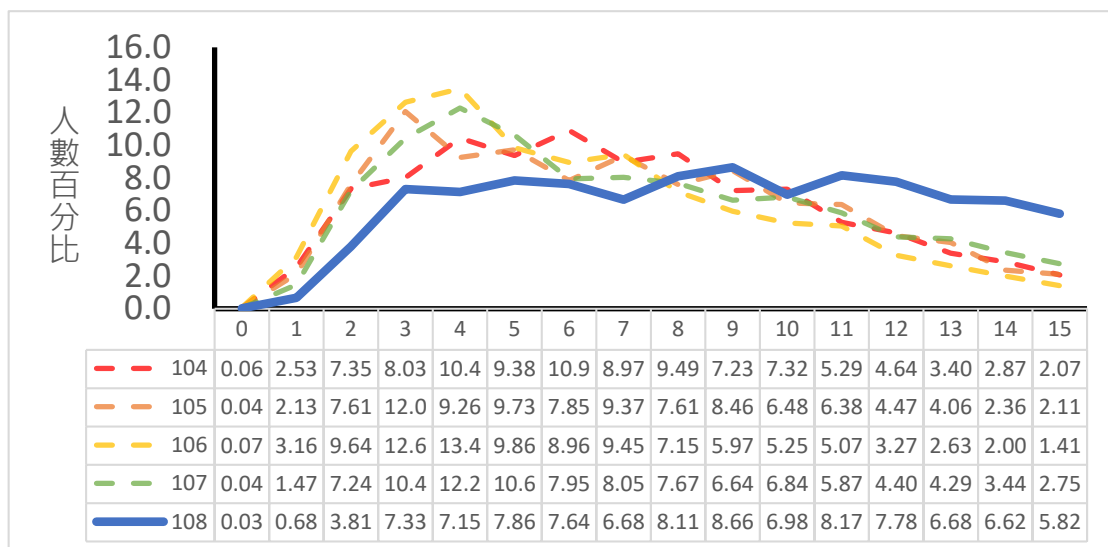


圖 1 近五年學測數學考科級分人數分布

六、若要特別區分高分群，組卷時需要配置較多偏難試題；若欲區分整體，組卷時就應適度配置各種難度試題。在考試時間有限且要以一份試卷評量所有考生之下，實難兼顧。組卷時若增加偏難的題目，可能可以拉出右尾，有助於區分高分群，但同時會導致過多人集中在左邊，等於把考生擠壓到較低得分組，此情形不僅減少考卷的整體鑑別度，可能還會有放棄學習的後遺症。

小結：

- (1) 觀察近五年學測數學科資料，108 年學測數學，不論是選擇題的平均鑑別度，或是整卷鑑別度(得分標準差)，不但沒有比過去四年差，反而是近五年最佳。
- (2) 社會議論的數學科「鑑別度」不足問題，與本中心數據所呈現的並不一致，大考中心必須特別向社會說明。

貳、穩定度

- 一、測驗的穩定度通常是指跨年比較時，各種參數的改變程度，目前社會對本次學測穩定度的議論，是屬於難易度部份，也就是級分人數分布的改變。
- 二、影響考生成績分布改變(不穩定)的因素很多，例如：考生能力的改變、學習內容的改變、試題內容設計改變(評量重點或層次的調整)、組卷規劃改變(題型、難易度等)、考試時間改變，以及其他非預期因素如命題不慎等。
- 三、以學測國文、英文、數學、社會、自然五考科近年的級分人數分布圖，可發現：
 1. 近五年來，國文、英文、自然、社會四科的分布型態(請見圖 2~圖 5)大致穩定，但略有起伏。
 2. 以數學來看，今年數學的圖形變動相對較大(請見圖 1)，左峰下降(低分組的區分功能提升)，相對而言，右尾的人數增多(高分組區分功能下降)。

四、單純要求穩定（級分人數分布不改變），相當於將過去的分佈視為固定的標準，不容改變，如此可能限制命題精進，也有利弊取舍的問題。

小結：

影響試題難度穩定度的因素很多，如何提升穩定度是大考中心須持續面對的挑戰。

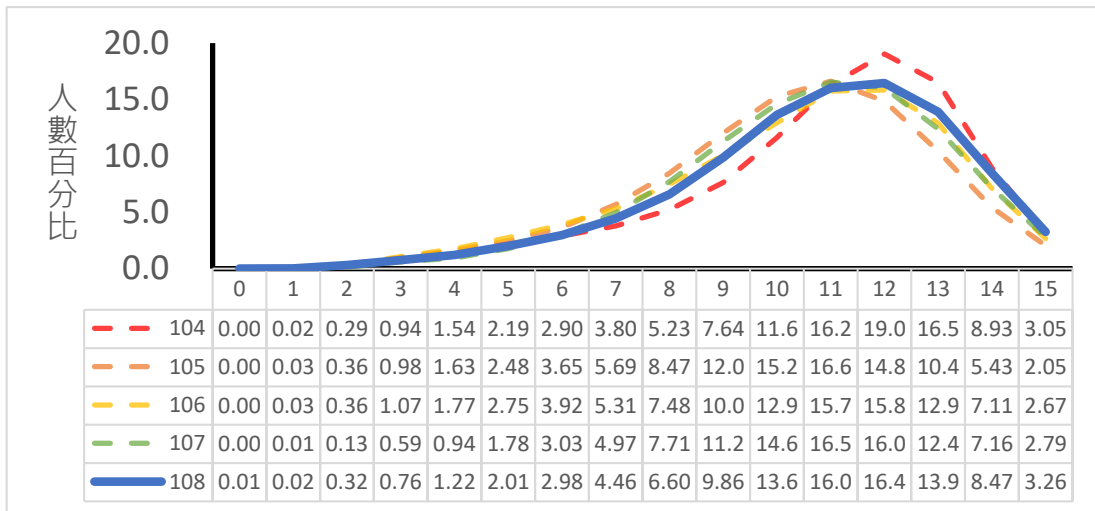


圖 2 近五年學測國文考科級分人數分布

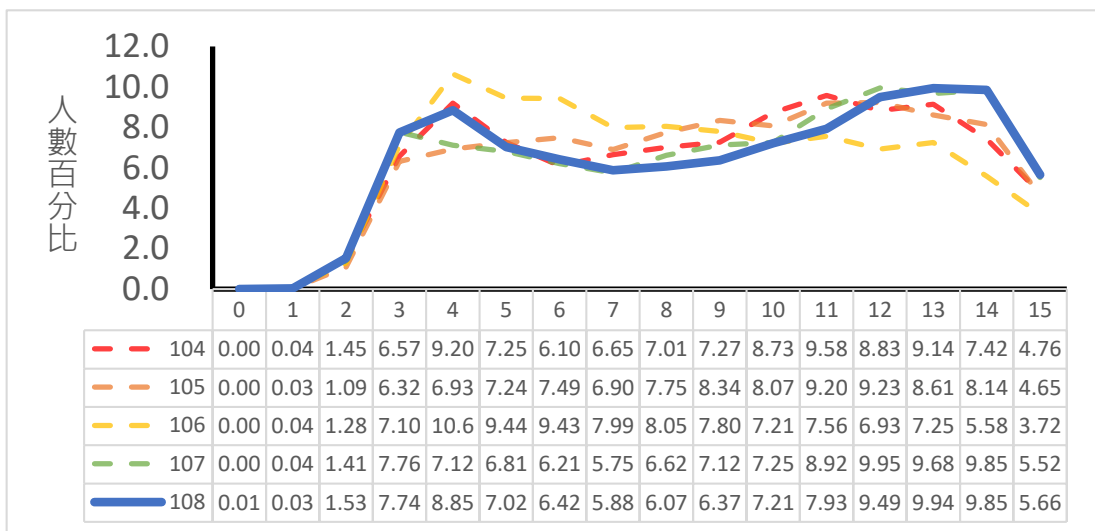


圖 3 近五年學測英文考科級分人數分布

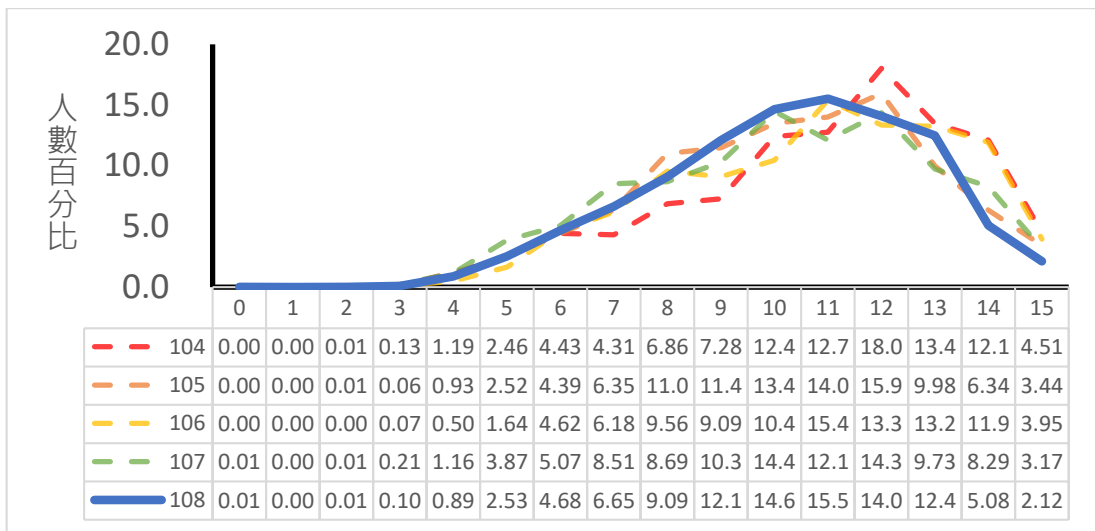


圖 4 近五年學測社會考科級分人數分布

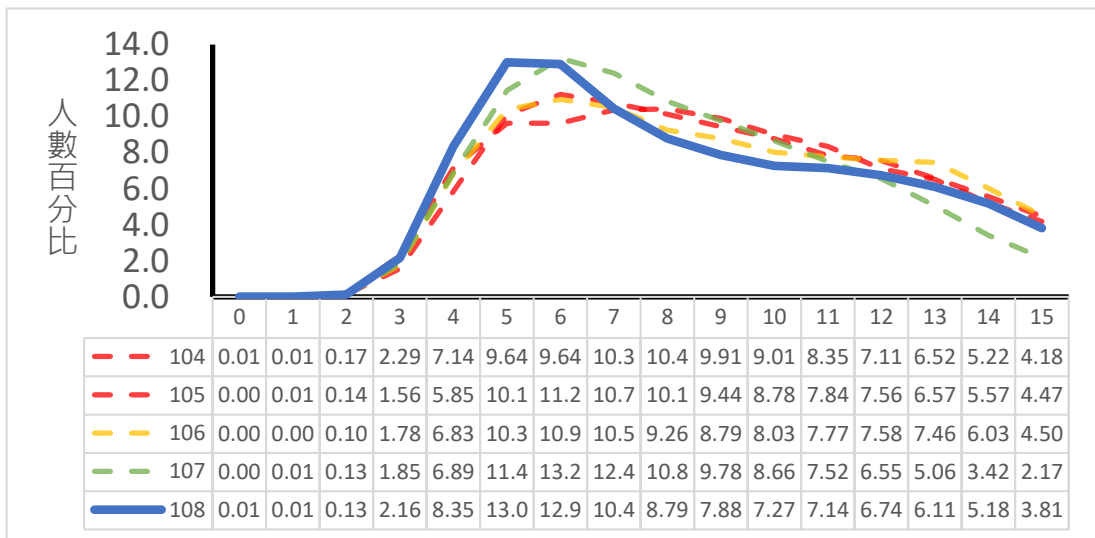


圖 5 近五年學測自然考科級分人數分布

參、級分

- 一、學測於民國 83 年首次辦理，成績以級分呈現，並用於申請第一階段之檢定或篩選。即使考生第一階段各科都獲得滿級分，仍需參加各校系辦理之第二階段指定項目甄試，因此，各校系需自訂不同的方式選擇學生，不會完全倚賴學測成績。
- 二、目前級分設計下，第 15 級分分數區間，會是所有級分分數區間中最大的。因此，即使錯部分試題也可獲 15 級分，考生可以不必執著於每一題都要答對。

三、在民國 83 年的首次學科能力測驗，級分數採用 10 級分。民國 88 年，配合大學推薦甄選入學招生回歸為大學正常招生管道，由 10 級分改為 15 級分。民國 91 年時，曾有大學建議修訂級分數，經討論，維持 15 級分。

四、回顧制度設計，學測使用較粗的級分，是希望各大學不要只憑分數篩選學生，能以更多元面向或資料評估學生能力與潛力。

五、級分數的增加可能減少申請第一階段的通過人數，但減少的幅度不確定，需要整合考、招資料進行模擬。而對於是否要改變級分數以減少第一階段通過人數（利弊），各大學校系、各高中、家長、考生及社會各界，看法未必一致。

六、增加級分數有很多可能作法，但是每種可能作法都有對應的後效，必須審慎評量。

小結：

大學多元入學方案將申請管道設計為二個階段，學測成績主要使用於第一階段之檢定或篩選，後續仍須搭配第二階段之校系指定項目甄試。級分數過去雖曾改變，但級分數的更動牽涉甚廣，若級分數增加，理論上或可以減少申請者的競爭性及不確定，但與申請入學管道與學測設計理念未必相符，且改變級分非大考中心權責範圍，決策屬於招生權責單位。