

以標準參照的入學考試 和班級評量促進科教發展

標準參照的評量方式是國際趨勢，台灣也正積極推動。本文說明其意義，以及標準參照的評量方法，如何影響升學考試、班級評量和診斷測驗。

宋曜廷

眾所周知，憲法是建設國家的基本藍圖。同樣的，在國民教育階段中，執行教育工作的最高指導原則，就是課程標準（curriculum standards），也就是台灣所稱的課程綱要和能力指標。

課綱：教育實務的最高指導原則

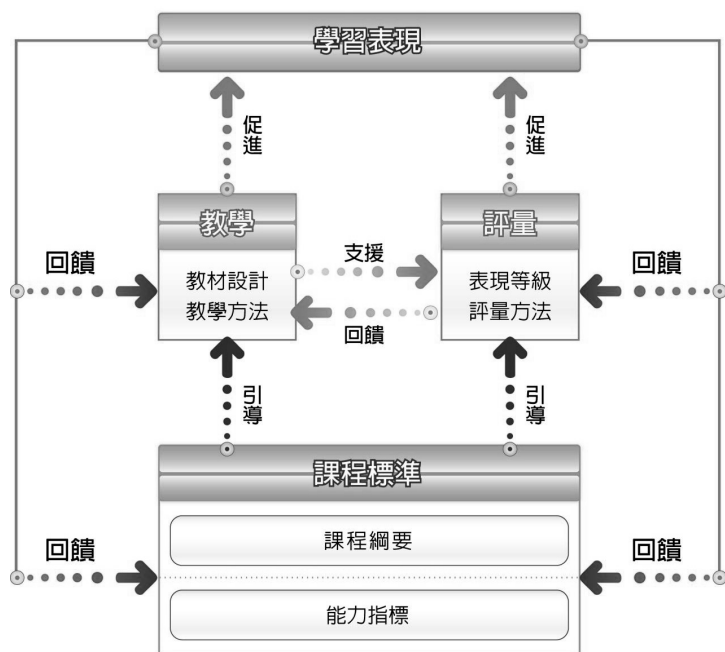
課程標準之所以重要，在於它明訂了一個國家的學生在每一個學習階段中，所應學習到的知識、技能、態度和行為等等。簡言之，課程標準描繪了我們希望培養的下一代在各方面的基本表現樣貌。以台灣的九年一貫課程綱要和能力指標為例，國中一年級學生在數學領域應該學習的內容有「數與量」和「代數」兩個主題，且「數與量」中的「因、倍數」部分，學生需要達到的能力表現之一為：「能『理解』質數的意義，並『認識』100以內的質數。」由上例可知，課程標準之所以為標準，因為它明訂了學習的方向與範疇，以及希望學生可以達到的能力。

課程綱要除了以學習範圍和能力表現作為明確的規範外，對於教師的教學和評量實務，也有重要的導引作用。例如九年一貫課綱中的「自然與生活科技領域」，指出國中學生應有的認知層次之能力有下列幾項，其一為「知道地球在宇宙中的相關地位」；其二為「知道簡單機械與熱機的工作原理，並

能列舉它們在生活中的應用」；其三為「在同類事件，但由不同來源的資料中，彙整出一通則性」。上述三種課綱揭示的能力，代表不同層次的教學目標，第一項為瞭解或記憶知識的層次；第二項為應用知識的層次；第三項為歸納和統整知識的層次。這三種教學目標必需靠不同的教學方法才能達成，如果只靠背誦和練習的教學方式，或許能讓學生將知識記牢，卻不可能達到活用知識的目的，也不可能引導學生覺察若干知識或發現之間的異同，養成歸納和分析的習慣。

因此，課綱中的能力指標，已經隱含了不同教學方法的意涵，也暗示並期待老師應該依據不同的能力設計不同的教學方法。同樣地，課綱所揭示的不同能力描述，也導引老師的評量方法。要瞭解學生對於知識的記憶或理解程度，多數時候一般的題型如是非題、選擇題和填充題可以達成。但在應用的層次，可能要加上讓學生實際操作的題型較能夠瞭解學生是否能將所學應用出來。至於要知道學生是否能對知識進行整合和批判，多數時候開放式的題型，如問答題或申論題，可能更能達到目標。

由上述說明可知，課程標準（或課綱）規範引導了教學和評量實務的進行，因此是學校教育實務的重要綱領。但經由老師在教學和評量的實作，以及學生的表現，也可以



圖一：課程標準、教學、評量和學習表現間的關係。

對課綱的適切性和可行性提供回饋和修正的作用。課綱、教學和評量三者之間的關係可以圖一來說明。

教學現場： 課綱領導？考試領導？

前已提到課綱、教學與評量是引導學生學習的三大主軸。課綱明確訂出學生在每一個特定階段應該要學習的內容和應獲得的能力；教學針對課綱明訂的知識技能進行培育；而評量則針對教學的效果和學習的成效進行評估，並作為回饋的依據。近二十年來，課綱有許多明確的改變，對於引領學生能力的發展方向也有諸多著墨，然而，相較於課綱的改變，學校教師在教學和評量上的變革，可以說極其有限。在2001年九年一貫課程實施後的頭幾年，政府已提倡合科課程與教學、多元評量等觀念，但這些觀念或作法在近幾年並沒有持續在教室中發揮活化教學和提升學習的效果，甚至多數不了了之。主要原因有三：

一、國內的升學主義盛行，教師多以考試的內容和題目形式馬首是瞻，跟升學考試（國中基測、大學學測）有關的科目，有考

有教，沒考不教。雖然國中基測的題目和取材已經盡量鼓勵活用，朝生活化和應用化的方向出題，但畢竟紙筆測驗選擇題的方式對於若干技能（如透過實驗培養的科學過程技能）和需要綜合評判的技能（如數學科的證明題），仍有很大限制，因此老師在此方面的著墨較少，學生亦不重視此方面的知識能力。

二、課綱本身對於教學和評量實務的引導不夠明確。雖然課綱中有揭櫫教學範圍和欲達成的能力，但似乎對於教科書編撰時的內容規範，效果遠大於對教學方法和評量方法所產生的影響。主要是因為課綱言簡意賅，僅提示內容和能力表現，對於教學方法並沒有提示，因此教師必須要根據內容和能力，思索可以達成教學目標的方法，這種隱晦的連結，讓老師無所適從，導致最後連課綱也不參考，便直接就教科書的內容，以講述的方式進行授課。

三、在評量方面，由於教科書商競爭激烈，因此設計許多教學輔助工具，評量光碟也因此盛行，許多教師可以直接自光碟中尋找合用的評量內容修改或援用，因此對課綱中所揭櫫的能力應該如何評量，多半一知半解。由於沒有深入理解和實踐課綱、教學和評量之間彼此相輔相成的關係，所以對多數教學現場的教師而言，課綱所揭櫫的能力、教師的教學、評量的方式，三者教學現場中往往是斷裂散軼而少有對應的（unaligned）。此種散軼的情形，起碼有三種影響：一、教師對於課綱中所期望達到的學生能力瞭解不深入，導致培養不同能力的方法並沒有受到重視，因此教學能力難以提升。二、升學考試，如大學學測或國中基測，成為引導教學的主要依據，過度練習變

成常態，學生學習的向度被窄化，許多高層的能力沒有被培養。不論老師、學生或家長由於缺乏對自身學習方向和內涵的信心，因此將國際學習評比（如PISA或TIMSS）視為引導國內評量或教學思考最常被提到的「標竿」作法。三、坊間參考書和測驗卷盛行，教師評量能力難以精進，加上對課綱的瞭解不深入，整體評量的方式主觀籠統，教師對學生進行評量的專業能力有下滑的隱憂。

教育現場中「教學」應該比「評量」對學生有更直接的培育功能，但目前國內的氛圍中，學校外部的入學考試和學校內部的班級評量，短期內仍是影響教學的主要力量。十二年國民基本教育推動在即，外部考試（如基測或會考）在高中職的入學功能將逐漸降低，而對學生學習表現的監控功能將逐漸增加。此外，班級評量如何確切發揮功能，達到維繫學生學力的目的，均是未來重要的措施。近期無論是高中新課綱或九年一貫課綱均有不少變動，此時正宜思考如何將課綱、教學、與評量進行貫串（alignment），以真正達到課綱領導教學和評量，評量和教學實現課綱目標的功能。本文將以標準參照評量（standard-based assessment）的觀點，以國中自然科為例，探討入學考試和班級評量應如何變革，逐步達到上述三者貫串並提升科學學習的目的。

標準參照評量

標準參照評量是將評量結果與事先預定的標準做比較，然後瞭解評量表現與所訂標準的差異。所謂標準可能是明確的分數水準，例如老師在學期初訂定出來的期末及格分數；公務員高考訂定多少分數達錄取水準；或是技術士檢定規定固定時間必須完成



宋曜廷在「2012年張昭鼎紀念研討會」演講中表示，標準參照評量應用在學力檢定將能發揮實質的功效。宋曜廷現為台灣師範大學心理與教育測驗研究發展中心主任，專長為教育心理學、心理與教育測驗等。

的術科項目有幾項；汽車駕照的筆試及格標準等等。除了量化的數字，標準也可能是對於某些學習目標的質性描述。例如學生在學習生物科一年後，他的表現等級可以如何劃分。以表一「國民中學學生學習成就評量標準」生物科中次主題「生物體內物質的運輸」為例。評量標準有幾個要素：「內容標準」，乃依據課綱所定義出來的學習內涵；「表現等級」，如表一的A、B、C、D、E；等級描述，如表一對各等級的各項表現標準之文字敘述。

根據表一的評量標準，可以達到幾個效果：一、引導教師更加重視課程標準。評量標準是依據課綱裡的教材內容和能力指標綜合整理，並且更具體的劃分出課綱中指出的學習目標，並加入課綱沒有明白劃分出來的能力表現等級，因此有助於教師更方便地將課綱融入教學目標當中。二、有效引導教師的教學。教師可具體知道教學內容中的哪些知識技能層級是被重視的，因此必須思考相對應的教學方法。三、評量結果客觀具體。傳統的百分制中老師將學生的成績分為0~100，約定俗成60分為及格。但及格所代表的意涵為何？生物科得到60分的學生學會了什麼？哪些沒有學會？而80分的又

如何？其說法多是籠統而主觀。若根據具體的表現標準評定成績等級，而非僅有一個分數（如 80 分），則可以知道其所對應的學習成果為何。四、指引學生明確學習地圖。評量標準所劃分的等級，將知識、技能或情意的學習目標明確的依等級陳述出來，能讓學生知道哪些表現內涵是受到鼓勵的，值得追尋的。

標準參照評量的作法在許多先進國家例如澳洲、英國、美國、新加坡、香港等都已經採用多年。這些國家一方面將標準參照評量應用在學力監控上，做為其設計會考的依據，另一方面也將評量標準的概念引進教師日常的班級評量當中，藉以改進教師的評量專業，進而促進學生的學習效果。以下將分別說明標準參照評量在入學考試、班級評量和診斷測驗的應用及其影響。

標準參照評量在入學考試的應用

標準參照評量若應用於離校考試，主要目的為學力監控，讓學校和主管機關瞭解學生學力的變化。例如，預計 103 年進行的國中教育會考，將自然科的表現標準分為三級，如表二。等級的描述有助於引導更多高層次試題的出現。以往國中基測因為題型的限制，要透過選擇題來測量高層次思考較為困難，所以 PISA 評比中的「辨識科學議題」、「科學地解釋現象」等能力雖然基測都有類似

題目，但在「科學舉證」上，因為需要利用開放式題目來促使學生提出科學證據以支持本身的主張或論點，所以目前基測難以達成（以選擇題為主）。但未來會考題型改變，就更容易測量到學生融會各概念的能力。

以數學科為例，會考將加考非選擇題型，評量國中生的數學溝通能力，以更全面地了解學生的數學能力表現，同時也有助學校教師的教學。即使高中職招生端有意將會考的等級作為入學的參考，僅三等級的劃分也有助於降低學生分分計較的壓力，讓學生的學習和教師的教學可以從反覆練習和死背硬記中解放出來，朝高層次的學習邁進。

關於自然科中何種題目的設計方式可以知道學生是否達到精熟的程度，還是僅達到

表一：國民中學學生學習成就評量標準——生物科等級描述

內容標準	主題	生物的構造與功能
	次主題	生物體內物質的運輸
表現等級	A	1. 能合理延伸植物體內物質運輸及蒸散作用的相關概念。 2. 能合理延伸人體循環系統運作的相關概念。 3. 能依指示完成植物體內水分運輸的相關實驗並能合理延伸。 4. 能依指示完成血液循環的相關實驗並能合理延伸。
	B	1. 能熟悉植物體內物質運輸及蒸散作用的過程。 2. 能熟悉人體循環系統的運作。 3. 能依指示完成植物體內水分運輸的相關實驗，並能提出合理的結論。 4. 能依指示完成血液循環的相關實驗，並能提出合理的結論。
	C	1. 能知道植物體內物質運輸及蒸散作用的構造與功能。 2. 能知道人體或動物的循環構造與功能。 3. 能依指示完成植物體內水分運輸的相關實驗，並確實記錄結果與透過引導提出結論。 4. 能依指示完成血液循環的相關實驗，並確實記錄結果與透過引導提出結論。
	D	1. 僅能部分知道植物體內物質運輸及蒸散作用的構造與功能。 2. 僅能部分知道人體或動物的循環構造與功能。 3. 僅能依指示部分完成植物體內水分運輸的相關實驗，並記錄結果。 4. 僅能依指示部分完成血液循環實驗，並記錄結果。
	E	未達D級

表二：教育會考自然科「精熟」、「基礎」及「待加強」3個等級描述

等級 考試科目	精熟	基礎	待加強
自然	能融會貫通學習內容，並能運用所培養的能力來解決需要多層次思考的問題。	能知道及理解學習內容，並能運用所培養的能力來解決基本的問題。	能部分知道及理解學習內容。

基礎或待加強的程度，我們可以透過國中基測的兩個題目來說明。

例一為評量學生是否能理解電路圖及並聯電路中的電壓與電流關係，作答此題時，學生須先判斷開關接通前後電路的差異，再利用三個電阻的電阻值相同及其在電路中並聯的關係，進而推算出通過安培計的電流。此試題屬於多層次思考的問題，學生須融會貫通學習內容，並運用所培養的能力才能解決此類型的試題。答對此類型試題的學生，其能力被歸為「精熟」等級的機率較高。

例二為評量學生是否能熟悉植物光合作用的過程。此題的題幹不同於學生的教材內容，但情境敘述較簡單。學生若能知道及理解植物進行光合作用時會消耗 CO₂，即可推測出正確答案。此試題屬於基本的問題，

學生僅需知道及理解學習內容，並運用所培養的基本能力就能解決此類型的試題。答對此類型試題的學生，其能力被歸為「基礎」等級以上的機率較高。

標準參照評量 在班級的應用

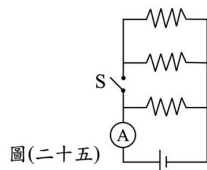
如果把前述評量標準的概念放到班級中應用，即「標準參照班級評量」(Standard-Based Classroom Assessment, SBCA)，指透過教師設計課程、評量和給分的過程，將標準參照的概念融入班級課程中，並與國家的評量標準相呼應的整體流程。

把評量標準融入班級評量，更能發揮引導教學和學習的效果，原因是在離校考試或入學考試中往往因為遷就紙筆測驗，而有許

多考試題型的限制，讓原本標準中的高層次表現無法適當展現。在班級評量中，則可以透過各種不同的形式以評量學生各類的能力，例如未來自然科的評量內容須配合自然科評量標準的目標，所有能力不可能只透過單一的評量方式，如評量「實驗程序知識」，可使用紙筆測驗；評量「設計實驗技巧」，可使用檢核表的紀錄；評量「作實驗的技巧」，可使用選取器材、安排布置器材和做實驗過程的實作評量。結合多元評量和多次評量的成果，

電路裝置如圖(二十五)所示，通過安培計的電流為1 A。若三個電阻器的電阻均為R 歐姆，且均符合歐姆定律，導線、安培計及電池的電阻可忽略不計，則在接通開關S 後，通過安培計的電流應為下列何者？

- (A) $\frac{1}{3}$ A (B) 1 A
(C) 2 A (D) 3 A (正解)



圖(二十五)

例一

小畢把池水、水草和魚裝入透明玻璃瓶，測量pH值後將瓶口密封，如圖(四)所示。將此瓶持續照光2天後，瓶內的生物仍生長良好，但測量得知瓶內水的pH值降低。在持續照光的條件下，若小畢想使瓶內水的pH值回復到接近照光前的數值，則可採用下列哪一方法？

- (A) 再多加一些魚在瓶內
(B) 將CO₂灌入瓶內水中
(C) 再多加一些水草在瓶內 (正解)
(D) 將瓶內的池水倒掉一半



圖(四)

例二

能夠得知各種面向的學生表現，並且能將評量的結果回饋教學以改進教學的品質，而非透過幾次總結性評量就決定了學生的學習成就。即便是採用紙筆測驗，標準參照班級評量也比較容易在期中考或段考中加入開放式題型，讓學生的高層思考可以展現。

下面的題目即為標準參照班級評量紙筆測驗的例子。例三的評量內容為生物科的分類與生態——生命的多樣性，評量目標為學生是否能理解生物五界（此說法較常使用）的分類與特徵，並且能利用生物特徵製作檢索表。題幹中給予動物簡易分類檢索表的作法，若學生具備歸納與分析的能力，即可利用植物特徵製作出新的檢索表。圖二、圖三和圖四為學生在此評量作業上的A、B和C等級表現樣卷示例（得到A、B和C等級的答案尚有其他可能，內容和評分規準因篇幅所限省略）。



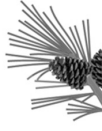

標準參照評量 在診斷和補救的應用

評量的目的的一方面是了解學生表現有多好，或者說引導教師或學生注意高層次能力的培養。前述內容多著墨於評量標準對於高層能力的可能影響。評量的更積極目的，是了解學生表現有多不佳，有多少學生在特定知識領域是沒有達到最基本的及格門檻，因此需要進一步的協助。

事實上評量標準中對於表現等級中下的描述，如表一中的D等級或表二的待加強等級，也可作為學習困難的學生診斷和補救的指標，發揮重要功能。由於表現等級有具體的表現描述，屬D等級或E等級

的學生將可以被明確地定義出來，進行及時的補救教學。由於學科知識有縱向的連貫性，因此學生在特定年級中所欠缺的知識，有可能是在更低年級就不會了。例如高二的學生在學習力學上如果學習效果不彰，可能是他在國三的時候在學習力學時就有困難。如果未來能將小學、國中、高中的評量標準串連起來，則能更有效追蹤學生學習困難是否發生在更前端，確切的問題點是在那一年級的哪些知識點，如此可以進行更有效的診斷和補救。

下圖為四種不同生物局部外形示意圖及其特徵。試回答下列相關問題：

甲	乙	丙	丁
			
有花朵	有孢子囊	有毬果	有菌絲

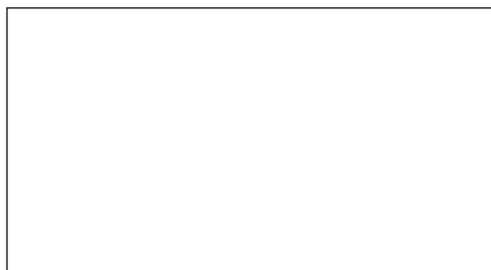
(1) 將上述四種生物分類，用甲、乙、丙、丁代號寫在下列作答處。
 屬於真菌的有：..... 屬於裸子植物的有：.....
 屬於蕨類的有：..... 屬於被子植物的有：.....

(2) 上述四種生物除了細胞膜、細胞質外，還有哪些共同的細胞構造？（寫出兩項即可。）

(3) 若將甲、乙、丙分為一組，丁為另一組，除了用菌絲的有無作為分類依據外，還可用哪一特徵作為分類依據？（寫出一項即可。）

(4) 參考下列動物簡易分類檢索表的作法，為上述生物甲、乙、丙、丁繪製一簡易檢索表。過程中「避免」以花朵、孢子囊、毬果、菌絲的有無作為分類依據，試以其它特徵的有無來區分並繪於下方方框處。

動物分類	代號
爬蟲類	戊
棘皮動物	己
兩生類	庚
節肢動物	辛



例三

結語

課綱、教學、評量三者彼此間散軼不對應，是多年來教學現場普遍的現象。這種對三者關連性的輕忽，也導致了在教學與學習上的諸多問題。本文以自然科為例，倡議透過標準參照評量，讓課綱、教學、評量三者可以適當貫串，並且有效引導入學考試、班級評量和診斷與補救的發展，進而引導教學的改變。

十二年國教實施在即，如何就由評量方法的改變，讓學生的基本學力可以維繫，而且高層次能力有所發展，是為最重要的目標。標準參照評量是一個重要的起步，當教師能夠對於課綱、教學、評量三者的內涵和關連了然於胸，代表他們對於國家知識工程的藍圖能夠融會貫通，則振興學生能力，指日可待。

宋曜廷：任職台灣師範大學
心理與教育測驗研究發展中心

圖二：A等級表現示例。

(1) 將上述四種生物分類，用甲、乙、丙、丁代號寫在下列作答處。
屬於真菌的有：丁 屬於裸子植物的有：丙
屬於蕨類的有：乙 屬於被子植物的有：甲

(2) 上述四種生物除了細胞膜、細胞質外，還有哪些共同的細胞構造？(寫出兩項即可。)
細胞核、液胞

(3) 若將甲、乙、丙分為一組，丁為另一組，除了用菌絲的有無作為分類依據外，還可用哪一特徵作為分類依據？(寫出一項即可。)
有無葉片

(4) 參考下列動物簡易分類檢索表的作法，為上述生物甲、乙、丙、丁繪製一簡易檢索表。過程中「避免」以花朵、孢子囊、莢果、菌絲的有無作為分類依據，試以其它特徵的有無來區分並繪於下方框處。

動物分類	代號
爬蟲類	戊
棘皮動物	己
兩生類	庚
節肢動物	辛

樣卷說明：

- Q1、學生能分辨出甲、乙、丙、丁分屬為被子植物、蕨類、裸子植物、真菌。
- Q2、學生能寫出甲、乙、丙、丁細胞的兩項共同構造。
- Q3、學生能正確地比較甲乙丙與丁之間不同的特徵。
- Q4、學生能根據生物特徵製作合理的檢索表。

圖三：B等級表現示例。

(1) 將上述四種生物分類，用甲、乙、丙、丁代號寫在下列作答處。
屬於真菌的有：丁 屬於裸子植物的有：丙
屬於蕨類的有：乙 屬於被子植物的有：甲

(2) 上述四種生物除了細胞膜、細胞質外，還有哪些共同的細胞構造？(寫出兩項即可。)
A: 液胞、粒線體

(3) 若將甲、乙、丙分為一組，丁為另一組，除了用菌絲的有無作為分類依據外，還可用哪一特徵作為分類依據？(寫出一項即可。)
A: 有無葉片

(4) 參考下列動物簡易分類檢索表的作法，為上述生物甲、乙、丙、丁繪製一簡易檢索表。過程中「避免」以花朵、孢子囊、莢果、菌絲的有無作為分類依據，試以其它特徵的有無來區分並繪於下方框處。

動物分類	代號
爬蟲類	戊
棘皮動物	己
兩生類	庚
節肢動物	辛

樣卷說明：

- Q1、學生能分辨出甲、乙、丙、丁分屬為被子植物、蕨類、裸子植物、真菌。
- Q2、學生能寫出甲、乙、丙、丁細胞的兩項共同構造。
- Q3、學生能正確地比較甲乙丙與丁之間不同的特徵。
- Q4、學生能根據生物特徵分類，製作的檢索表雖然將生物合理區分出來，但是使用的特徵（花朵）與題目中提出的特徵重複。

圖四：C等級表現示例。

(1) 將上述四種生物分類，用甲、乙、丙、丁代號寫在下列作答處。
屬於真菌的有：丁 屬於裸子植物的有：丙
屬於蕨類的有：乙 屬於被子植物的有：甲

(2) 上述四種生物除了細胞膜、細胞質外，還有哪些共同的細胞構造？(寫出兩項即可。)
細胞壁、細胞核

(3) 若將甲、乙、丙分為一組，丁為另一組，除了用菌絲的有無作為分類依據外，還可用哪一特徵作為分類依據？(寫出一項即可。)
葉綠體

(4) 參考下列動物簡易分類檢索表的作法，為上述生物甲、乙、丙、丁繪製一簡易檢索表。過程中「避免」以花朵、孢子囊、莢果、菌絲的有無作為分類依據，試以其它特徵的有無來區分並繪於下方框處。

動物分類	代號
爬蟲類	戊
棘皮動物	己
兩生類	庚
節肢動物	辛

樣卷說明：

- Q1、學生能分辨出甲、乙、丙、丁分屬為被子植物、蕨類、裸子植物、真菌。
- Q2、學生能寫出甲、乙、丙、丁細胞的兩項共同構造。
- Q3、學生能正確地比較甲乙丙與丁之間不同的特徵。
- Q4、學生僅能利用特徵將甲及乙丙區分，但是使用不適合的特徵(是否具有細胞壁)進行甲乙丙及丁的分類。