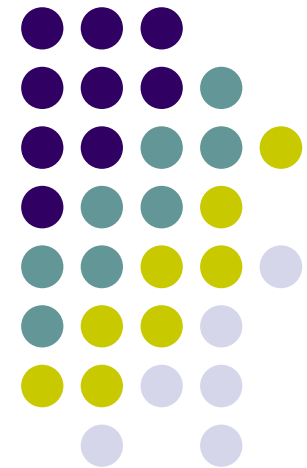
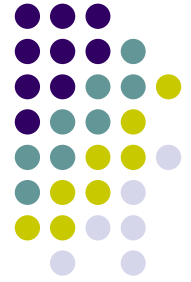


Chapter 02

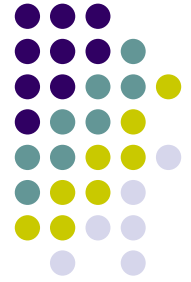
形成性評量與總結性評量





前言

- 教師與學生是教育過程中進行互動的主體，課程、教材是互動的內容
- 教學與評量則是主要的互動形式(歐滄和，民91)
- 評量扮演回饋的功能，用來瞭解學生學習成果及調整或擬訂教師的教學策略，以達到最佳的教學效果 (歐滄和，民91)



前言(cont.)

- 教學評量想要瞭解學生：
 - 學習前的起點行為在哪裡？
 - 學習中應採取什麼教學策略？
 - 學習時有什麼障礙亟待克服？
 - 學習後是否達成預期的教學目標？



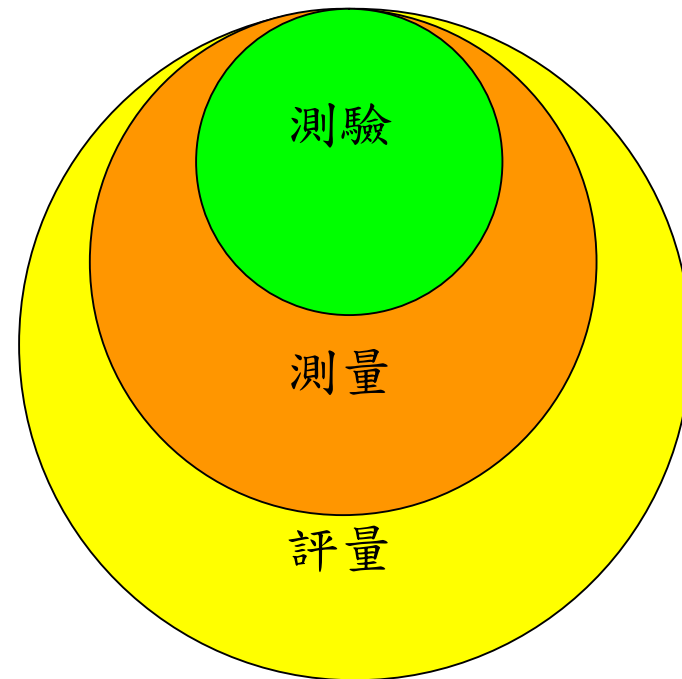
前言(cont.)

- 評量為教學的一項後續(**follow-up**)活動
- 學生及家長通常不瞭解評量的真正目的及評量結果所代表的意義，而僅強調其中所呈現的片面資料，例如成績、排名
- 錯誤的評量角色可能易客為主，對教學產生導引作用(**washback**)，亦即所謂的「考試領導教學」(施玉惠等，民88)

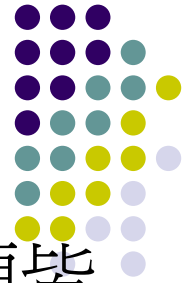


概念分析

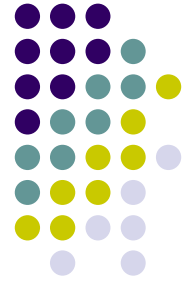
- 測驗(Testing)
- 測量(Measurement)
- 評量(Assessment)



測驗(Testing)



- 測驗(Testing)的範圍最小、也最明確，其試題皆有「正確」或「最恰當」的答案，如段考、隨堂考等。
- 本質上是一種觀察的工具(歐滄和，民91)，用以衡量受測者的某些行為樣本，進而推論其具有的心理特質或能力」。
 - 例如：期中考、期末考。
- 通常測驗的「成績」會被用來當作學生的學習成就。而實際上對學生而言，成績除了反映學生的學習成就之外，並無法協助瞭解個人的學習困難及改善的方向(黃國禎，民93)



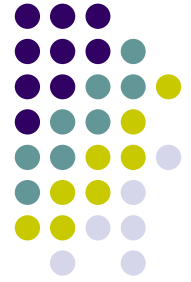
測量(Measurement)

- 所謂測量，就是指根據量尺，用數字來描述個人特質的歷程。
- 測量在本質上只是一種將事務的特徵加以數量化的過程，它並不包含質的描述，而且對於獲得的結果也不做任何的價值判斷。(歐滄和，民91)

評量(Assessment)的定義



- 評量(Assessment)的範圍較廣，以測驗、面談(interview)、專案實作 (project)、檔案評量(portfolio assessment)等方式，以診斷學生的能力。
- 是指從多種角度以多種方法去評估一個人的知識或能力，以作為教學或輔導效果的證明及依據。
- 在使用「評量」這個詞時，除了評估學生的能力之外，還包括協助學生解決問題。



評量的目的

- 瞭解學生的潛能與學習成就，作為判斷學生努力的程度。
- 診斷學生的學習困難與障礙，作為補救教學及個別輔導的依據。
- 估量教師教學的效率，提供教師改進教材教法的參考。
- 瞭解學生學習進步的情形，觸發學生的學習興趣。
- 提供家長參考，作為瞭解學生在校學習的情形。

Bloom et al. (1981) 對評量的看法



- 評量是一種用以確定學生學習水準和教學有效性的證據的方法。
- 評量包括了比一般期末考試更多種類的證據。
- 評量是澄清(**clarify**)教育的重要目的與教學目標的一種輔助手段，以及確定學生按這些理想的方式發展到何種程度的一種過程。
- 評量作為一種回饋-校正系統，在教學過程的每一步驟判斷該過程是否有效；如果無效，必須及時採取什麼變革，以確保過程的有效性。
- 評量是教育研究與實踐的一種工具，用於查明在達成某一教育目的時，不同的歷程是否同樣有效。



其他學者對評量的看法

- 評量是運用科學方法和技術，蒐集有關學生學習行爲及其成就的正確資料，再根據教學目標，就學生學習表現的情形，予以分析、研究和評斷的一系列工作(簡茂發，民88)
- 評量係指教師蒐集、綜合、解釋有關學生的各種資料，以協助教師進行各種「教」與「學」決定的歷程（李坤崇，1999）



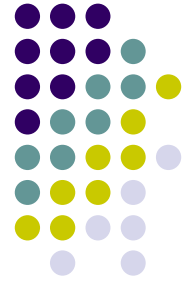
評量時機

- 依照教學前、教學中、教學後不同實施的時間點，可分為：
 - 安置性評量
 - 形成性評量
 - 診斷性評量
 - 總結性評量

安置性評量

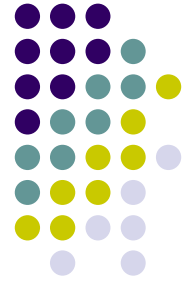


- 決定班級教學的起點
- 決定是否先行複習舊教材內容
- 決定選擇何種適當的教材和教法
- 決定如何將學生分組或安排在特殊班級中學習。
 - 依據學生的學習表現與需求，評估特殊性向與能力，提供適切安置。
 - 資賦優異甄選、未足齡兒童提早入學甄選、縮短修業年限、單科或全科跳級、身心障礙鑑定及教育安置。



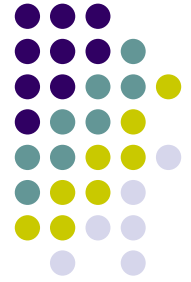
形成性評量

- 不斷提供回饋給學生和教師，使他們得知教學和學習的成功與失敗。
- 有助於教師了解學生在哪些方面的學習尚未達到教學目標所要求的程度。
- 評量的範圍較小，內容僅限於教學的特定內容，可能只是一個概念或原則。



診斷性評量

- 目的在診斷學生的困難所在，並針對其困難，予以必要的補救教學。
- 診斷性評量是一種更綜合性和精密性的評量。
- 學習困難特別嚴重者，需要進一步的去探討其真正的原因，有時需要使用不同的診斷工具，或轉嫁尋求專業人員的協助。



總結性評量

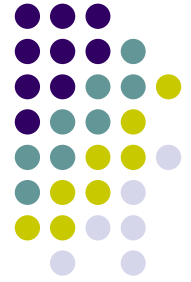
- 主要在評估本教學單元所列教學目標達成的程度，以及檢討所用的教學方法是否適當有效，並且評定學生的學習成果。
- 評量的測驗試題所涵蓋的難度範圍較廣，通常是抽取能代表學習內容的樣本作爲試題，且多爲常模參照。

總結性評量與形成性評量的比較



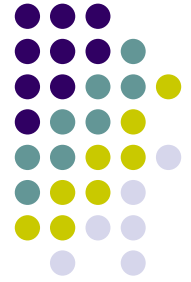
	形成性評量	總結性評量
一般名稱	課堂作業練習、單元測驗、精熟測驗	期中考、期末考、畢業資格考、標準化成就測驗
使用目的	1. 提供學生立即回饋 2. 作為矯正學習的依據	1. 評定學生學期成績 2. 評定教學方案或課程的有效性
試題取樣	以每一單元的教材為範圍；範圍小，代表性高	以整個單元或整學期教材為範圍；範圍大，代表性低
施測頻率與作答時間	單元兩次，次數多，但作答時間短	每學期兩三次，次數少，但作答時間長
試題類化程度	強調基本的知識與能力，類化程度較低	強調各單元間知識能力的綜合應用，類化程度較高

※ 歐滄和(民91)教育測驗與評量。台北：心理，471)



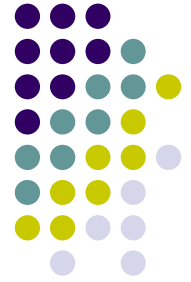
評量的方式

- 紙筆測驗
 - 學生經由教師依照教學目標、教材內容所編定的測驗。
- 實作評量
 - 實作評量是從做中學概念引申出來，藉由直接的現場觀察與評斷，或間接的從學生的作品去評量學習成效。
 - 例如：科學實驗、教學解題、寫作、口頭報告、戲劇、演出、闖關、小書製作



評量的方式(cont.)

- 動態評量
 - 動態評量是師生交互作用的過程，其目的在協助學生學習。
 - 教師以「測驗－介入－再測驗」的形式，對兒童的一般認知能力或特定學科領域進行持續性學習歷程的評量。
- 卷宗評量 (portfolio assessment)
 - 檔案評量源自於學生自傳文本的概念，其功能透過檔案歷程的記錄，了解並協助學生的成長。
 - 在有目的的蒐集學生的作品，依作品的品質與進步的情形做評量。



網路學習評量系統

- 應用於e-Learning 環境中，讓電腦協助教師與學生進行回饋互動的一個工具。
- 國內外學者依形成性評量策略所發展出來之e-Learning 輔助教學之工具有：
 - **PsyCAL** (Psychology of Computer Assisted Learning)
 - **Self-Assessment**自我評量模組
 - **WATA** (Web-based Assessment and Test Analysis System)網路評量與試後分析系統

PsyCAL (Psychology of Computer Assisted Learning)



- 國外學者Buchanan於1998年針對大學心理學課程所發展。
- 主要特徵是「重複作答」與「不提供答案」。
 - Buchanan 指出「重複作答」與「不提供答案」兩功能結合在一起，能激發學習者反覆的記憶。
 - 當學生作答錯誤時，會有提示性的即時回饋，導引學生找出正確答案；該策略可讓學生在練習與主動發覺答案的過程中，獲得更為精熟的學習，進而提升其學習效益。



Self-Assessment

- Gardner, Sheridan & White 在2002 年設計的網路形成性評量工具。
- 具有「重複作答」、「作答完後提供說明資料」、「發問」與「查詢他人成績」等功能
- 透過獲得即時的回饋，以提升學生的學習效益。

WATA (Web-based Assessment and Test Analysis System)

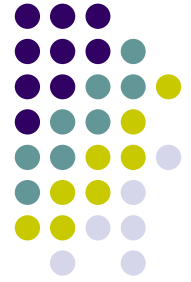


- 彰師大生物學系研發，包含總結性評量模組與形成性評量模組。
- 總結性評量模組主要用以瞭解網路學習的效益。
- 形成性評量具有「三次作答」、「發問功能」、「過關後查詢成績」、「過關後可查詢個人學習歷程」與「過關動畫」等設計。

網路形成性評量在效益上的研究



- **PsyCAL** 於大學心理學課程研究發現，使用**PsyCAL** 在學習效益上明顯高於未使用的學生(Buchanan, 1998)。
- **CECIL** 的「**Self-Assessment**」設計對網路學習效益有顯著的幫助，學生喜歡以此功能輔助自己在網路上進行學習。（ Gardner, Sheridan & White (2002) ）
- **WATA** 系統已實際應用於大學普通生物學的教學及推廣國中自然生活與科技教學上，發現使用**WATA** 形成性評量模組輔助學習，其學習效益越高。
- **WATA** 的「三次作答」、「過關後查詢成績」、「過關後可查詢個人學習歷程」與「**Flash** 過關動畫」等設計，在引發學習者的學習動機上也佔有相當重要的角色。



網路形成性評量策略

- 重複作答
 - 指定的題庫範圍中亂數隨機選取五題，題序與選項順序都採亂數排列。
 - 同一試題連續答對三次後，該試題即不再出現於往後的測驗中。
 - 同一試題無法連續三次答對，而在重複作答期間發生答錯現象，則該題答對次數的累積將會歸零重新計算。



網路形成性評量策略(cont.)

- 不提供答案：
 - 學習者送出試卷後，立即給予得分與呈現各個試題所作答的情況，但不給予正確答案。
 - 即時回饋的功能，由系統主動提供學習者「作答參考資料」，協助學習者自行找出答案。
- 發問功能：
 - 提供與教師非同步互動的機會，讓學生可以線上發問，與教師進行非同步互動。



網路形成性評量策略(cont.)

- 查詢成績

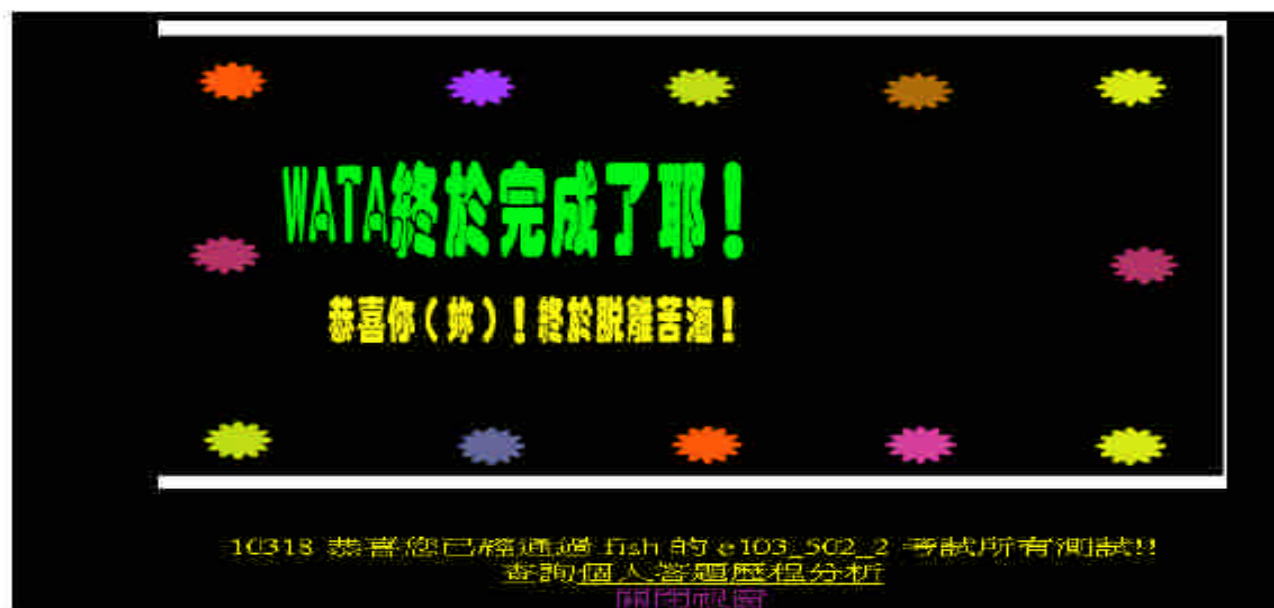
列出詳細列表

考試者	成績	考試時間	作答時間(分)
8923033 (第1次)	80	2003-12-07 19:20:08	1.4
8923041 (第1次)	100	2003-12-07 20:01:26	1.4
8923034 (第1次)	100	2003-12-07 20:11:44	1.37
8923008 (第1次)	100	2003-12-07 21:35:15	0.92
8923022 (第1次)	100	2003-12-07 22:01:22	1.25
8923038 (第1次)	80	2003-12-07 22:39:47	1.47
8923042 (第1次)	80	2003-12-07 23:14:18	2.89
8923009 (第1次)	80	2003-12-07 23:45:34	0.67
8923010 (第1次)	100	2003-12-08 00:16:49	1.39
8923011 (第1次)	80	2003-12-08 00:41:45	1.69
8923021 (第1次)	60	2003-12-08 16:16:32	0.87
8923013 (第1次)	100	2003-12-08 16:46:26	1.19
8923003 (第1次)	100	2003-12-08 18:22:49	1.22
8923023 (第1次)	100	2003-12-10 00:05:54	2.99
8923020 (第1次)	100	2003-12-10 23:04:16	1.49
8923035 (第1次)	80	2003-12-10 23:27:20	1.37
8923026 (第1次)	100	2003-12-11 22:05:18	2.27
8923002 (第1次)	80	2003-12-11 22:57:18	2.34
8923017 (第1次)	100	2003-12-11 23:26:00	1.87
8923032 (第1次)	100	2003-12-12 10:25:50	1.82
8923019 (第1次)	80	2003-12-12 18:00:22	1.54



網路形成性評量策略(cont.)

- 過關動畫
 - 藉由此新奇的視覺效果來引發其學習動機，並引發其完成測驗學習的成就感。

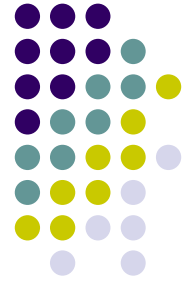




網路形成性評量策略(cont.)

- 過關後可查詢個人答題歷程

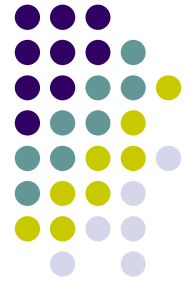
題目 全班:91.9%	一個穩定的生態系中，維持物質的循環，必須具備的條件是? 圖解:
選項	<ul style="list-style-type: none">Ⓐ (2.71%) 生產者與消費者Ⓑ (91.9%) 生產者與分解者Ⓒ (1.36%) 消費者與分解者Ⓓ (2.71%) 清除者與分解者Ⓔ (0%)Ⓕ (1.36%) [棄答]
編號	1114 生態系的生物組成
答案提示	生產者與分解者 南一版第六章生物與環境的交互作用
8923025 答對率: 80%	第 1 次: 生產者與分解者 YES!! 第 2 次: 生產者與消費者 第 3 次: 生產者與分解者 YES!! 第 4 次: 生產者與分解者 YES!! 第 5 次: 生產者與分解者 YES!! 積分:8



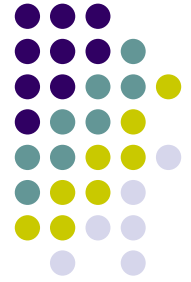
研究結論

- 重複作答策略：
 - 可以使學習者對學習內容更加熟悉，使學生可以在作題的過程中獲得成就感與警覺到自己的學習狀況。
- 不提供答案策略：
 - 可以提供機會使學習者更願意去把不了解的觀念加以澄清，而其中提供「作答參考說明資料」的設計，學習者表示可以因此獲得更多的學習機會。

研究結論(cont.)

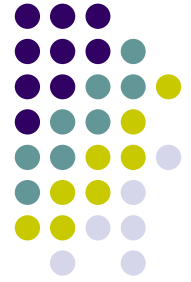


- 查詢成績策略：
 - 可以讓學習者瞭解他人的學習狀況，促進同儕間相互觀摩與良性競爭。
- 發問功能策略：
 - 可以提供學習者解決問題的機會外，更可以讓學習者在發問的過程中，瞭解自己不懂之處，以及對於與問題相關的觀念更加熟悉。



研究結論(cont.)

- 過關後可查詢個人答題歷程策略：
 - 可以提供學習者了解其學習狀況外，亦可使其警覺學習的缺失。
- 過關動畫：
 - 可以使測驗更加生動有趣，也可以激起挑戰過關與作題的意願。



結語

- 評量不是只看學生的缺點或弱處，而是要找出其學習優勢。
- 評量情境的設計應使各種不同能力的學習者都有成功的機會(應用多元智慧的理念)。
- 仔細觀察學習者細微的進步，並給予回饋與肯定。
- 不急著期待每位學生在同一個時間達到相同的學習目標，也不必期待所有學生在一次的教學中都完成所有的學習。
- 評量不一定需要採取個人間的競爭，合作學習的能力亦應給予適當的評量。



參考文獻

- William & Black (1996) 。 *closing the gap between actual and desired levels of performance*, p.543
- 邱淵等譯(民78) 教學評量。台北：五南。第一章。【譯自Bloom, B.S., Madaus, G. F. & Hastings, J. T. (1981) Evaluation to improve learning.】
- 歐滄和 (民91) 。教育測驗與評量。台北：心理出版社。7-9
- 李坤崇 (民87)：人性化、多元化教育評量—從開放教育談起。高雄市政府公教人力資源發展中心編印，多元教學評量， 91-133。
- 程俊博 (民93) 。以多元觀點評量生活科技。生活科技教育月刊，二〇〇四年三十七卷 第七期。
- 李坤崇 (1999) 。多元化教學評量。台北市：心理出版社。
- 簡茂發 (1999) 。多元化評量之理念與方法。教師天地，99 期，11~17 頁。
- 黃國禎 (2004) 。線上學習與評量系統的發展趨勢與挑戰。資策會數位學習技術中心，取自 http://www.elearn.org.tw/eLearn/eNews/eL_Insight_016.htm
- 王秀玲 (2006) 。課程理論的先驅：R.W.Tyler 《課程與教學的基本原理》。中等教育第57 卷第1 期2006 年2 月，131

參考文獻



- 林宜臻（2002）。形成性評量在數學學習領域課程設計之應用。台中市政府教育局國民教育輔導團電子報。No.0003，2002/12/10
- 林宜臻(2002)。讓促成思考過程之形成性評量在數學課堂生根。研習資訊，19(1)，8—16。
- 徐崇溫（民77）。結構主義與後結構主義。台北：谷風。
- 張明侃。教學評量與紙筆測驗命題技術探討。2007/05/30 取自 http://www.kuhes.tyc.edu.tw/~head/_private/study/18_test.doc
- 施玉惠、周中天、陳淑嬌、朱惠美、陳純音、葉錫南 (1999) 《九年一貫英語科教學與評量模式》第三章，教育部計畫編號881A153
- 教育部(民90)。國民中小學學生成績評量準則。
- 王子華、王國華、王瑋龍、黃世傑。不同形成性評量模式對國中生網路學習之效益評估。科學教育學刊。12 卷，4 期。(出版中)
- Buchanan, T. (1998). Using the World Wide Web for formative assessment. *Journal of Educational Technology Systems*, 27, 71-79.
- Buchanan, T. (2000). The efficacy of a World-Wide Web mediated formative assessment. *Journal of Computer Assisted Learning*, 16, 193-200.
- Gardner, L., Sheridan, D. & White, D. (2002). A Web-based Learning and Assessment System to Support Flexible Education, *Journal of Computer Assisted Learning*, 18, 125-136.