



學習成果導向評量： Rubrics 的運用與設計

李紋霞
教師發展組副組長
臺灣大學教學發展中心
leewsj@ntu.edu.tw

評量的目的

甄選性

甄選合適的人才

一定的標準明確的判準、名額的限制

預示性

推估預測(性向、偏好、才能)

常模當指標、理論當依據

工作性

瞭解現況以因應情勢使工作順利進行

教學評量屬於此類

班級教學之評量類別

安置評量 Placement

- 決定先備技能、課程目標精熟度及學習狀況
- 準備度測驗、自陳量表、性向測驗等

形成評量 Formative

- 決定學習進展、提供回饋增強學習或更正錯誤
- 自編測驗作業、觀察技術、出版社測驗

診斷評量 Diagnostic

- 決定持續性學習困難原因(心智、生理、情緒、環境)
- 標準或自編診斷測驗、觀察技術

總結評量 Summative

- 分派成績決定結束時的學習成就或認證目標達成的精熟度
- 自編測驗、表現評分量表、作業產品量表

高等教育評量發展趨勢

- 美國

- 績效責任 (accountability), 聯邦政府要求各區認證評鑑機構, 加強運用各校提供之評鑑結果, 作為重要認證指標。

- 歐盟

- 調整與對應 (tuning & alignment)
核心能力、課程、教學、評量均要相互對應

- 台灣

- 高教評鑑，著重「學生學習成效」

培養什麼樣的學生 (具備哪些核心能力)

系所訂定評量方法，以有效評量學生核心能力

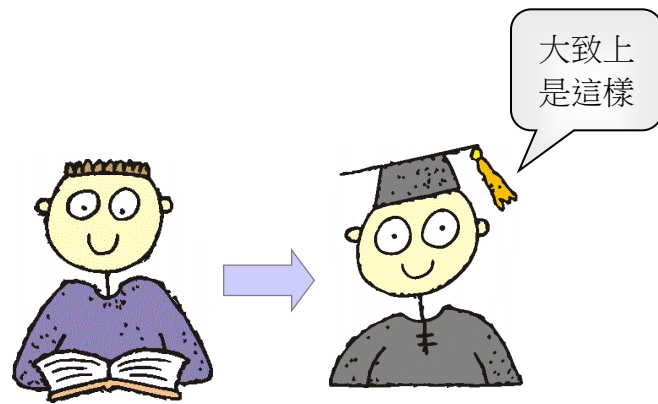
系所如何適當配置課程、經費及人力，使學生掌握核心能力

學生學習結果如何？

學生如無法達到目標, 校及系所的改善機制? 調整目標? 教法?
課程? 經費與人力配置?

就是 ...

學生經過學習之後所應具備的能力



核心能力應包含 **知識 (knowledge)**、**技能 (skill)** 與 **態度 (attitude)**，協助學生個人與未來職涯發展。

(Holmes & Hooper, 2000)

更詳細一點 ...

核心能力：(1) 是系所特色、(2) 讓外人瞭解系所核心價值、(3) 是系所的改善基礎以提升整體教育品質、(4) 讓學生明確知道學習目的，使學習更有重心、(5) 使教師之教學目標更清楚，藉此設計學習情境，並能明確評估教學效果。

訂定核心能力基本原則

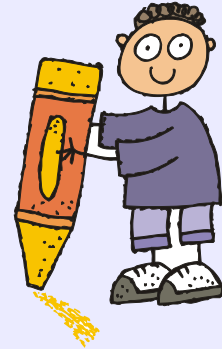
6

Adapted from 黃淑玲 (2010)
以學生學習成效為主軸的評鑑實務

基本原則：

- (1) 明確具體，但不狹隘。
- (2) 可被觀察，得以測量。
- (3) 以少為多，便利執行。
- (4) 教師參與，配合院校。

簡單明白
易懂易行



整體架構：

教育目標：
想教出怎樣的學生？

A	教育目標 A
B	教育目標 B
C	教育目標 C

核心能力：
如何達成教育目標？

1	核心能力 1
2	核心能力 2
3	核心能力 3
4	核心能力 4
5	核心能力 5

每個核心能力可經由選修幾門課程達成

100	課程
101	課程
200	課程
210	課程
311	課程
312	課程
400	課程
900	課程

課程設計

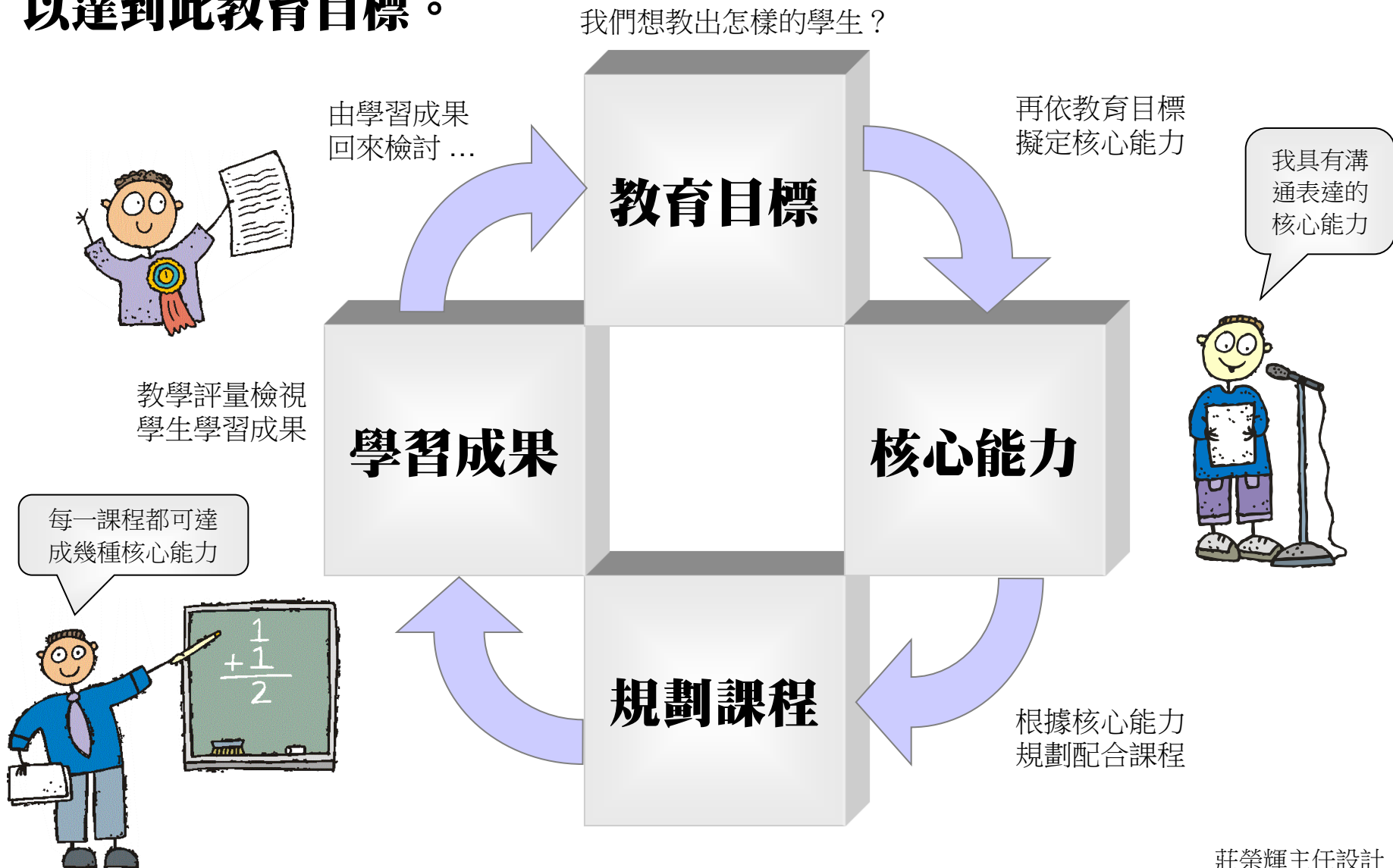
每個教育目標經由
數個核心能力達成

一門課程得以培養數種核心能力

如何訂定核心能力？

3

先凝聚系所『**教育目標**』為基本指南，再討論各項『**核心能力**』以達到此教育目標。



教育目標 4

Students should be proficient at using modern computer tools to solve problems.

Outcome 1 – Students will be able to develop customized solution software.

Outcome 2 – Student will know how to use high level modeling and computing tools such as spreadsheet programs, equation solvers, UML, and simulation software to analyze engineering problems.

藉由兩個核心能力 (Outcomes 1 & 2) 以達到 **教育目標 4**

訂定核心能力應考慮的關鍵問題

實施對象：

對畢業生的期望為何？

What is expected from a graduate of the program?

在學生之學習過程中，對他們的期望為何？

What is expected as the students progresses through the program?

關鍵內涵：

學生應該知道什麼？

What does the student **know**?

知識 (Knowledge)

學生可以做什麼？

What can the student **do**?

技能 (Skill)

學生應該在乎什麼？

What does the student **care about**?

態度及價值觀 (Attitude)

實例 1：伊利諾大學香檳分校 (比較與世界文學系)

<http://wasc.ucr.edu/docs/Sample%20Learning%20Outcomes%20--%20Humanities.pdf>

知 識

- (1) 廣泛瞭解至少兩種文學的歷史發展
- (2) 擁有分辨文本文化與歷史脈絡的能力

技 能

具備能夠清楚、一致、具說服力地表達自己的能力 (口頭與書面)

態度與價值觀

能夠賞析文本的美感

實例 2：舊金山州立大學 (化學及生化系)

8

核心能力	1	2	3	4	5	6	7	8	9
CHEM 340	★			★					
CHEM 343		★	★		★			★	★
CHEM 349	★								
CHEM 351	★			★					
CHEM 353	★			★					
CHEM 422		★	★		★			★	★
CHEM 425	★								
CHEM 426		★	★			★	★	★	★
評量方式	a, b	a, b	d	*	c, d	d	d	a, b	d

↑ (a) 考試表現 (b) 小組討論內容 (c) 書面報告 (d) 討論中

* 2007-08 運用 CHEM115、CHEM116 期末考題，評估修讀 CHEM300 或 CHEM301 前後之化學知識存留情形

← 核心能力 1~9：

1) 展現對重要概念的了解，並能解決下列五項化學次領域問題：分析化學、生物化學、有機化學、無機化學、物理化學。

2) 執行基本化學實驗程序，包括使用新的儀器進行合成、分離、離析、分析與詳述分子特質。

3) 闡釋化學與其他相關領域的關係，如生物科學、材料科學與環境科學。

4) 在進階課程中，展現對於先前所學知識的了解與整合(指利用多門課程的知識以了解新知)。

5) 蒐尋化學文獻撰寫報告，並閱讀與了解領域相關文獻。

6) 將課堂知識應用於實驗中，解決實驗問題。

7) 了解導出化學理論與模型的重要實驗。

8) 運用口語、書面與圖表等方式，有效傳達化學實驗結果。

9) 在小組中，對於解決化學問題有所貢獻。

教師層級評量實例：教育社會學

他評

1 專業知識

- Open book exam: 自己出題，自己回答
- 比較新舊英文期刊論文

2 溝通表達

- 書面報告：個人報告
- 口頭：小組報告

3 團隊合作

- 同儕評量：參與度及貢獻度 (避免 free rider) (1-10 級)
- 教師評量：報告的架構及邏輯性，看是否分工且合作 (3-4 級)

自評

1 專業知識

- 修完這門課後，我能掌握這門課重要內容與概念
- 這門課幫助我發現問題的能力

2 溝通表達

- 修完這門課後，我比較有膽量在公開場合作報告
- 這門課幫助我清楚表達自己的想法，使他人聽懂

3 團隊合作

- 參與小組合作後，我更知道如何善盡自己的職責
- 我學會小組報告時與組員既分工又合作

評分

公平公正地對待每個學生

並給予客觀的分數

反應學生的學習成效

進而引發學習的動力

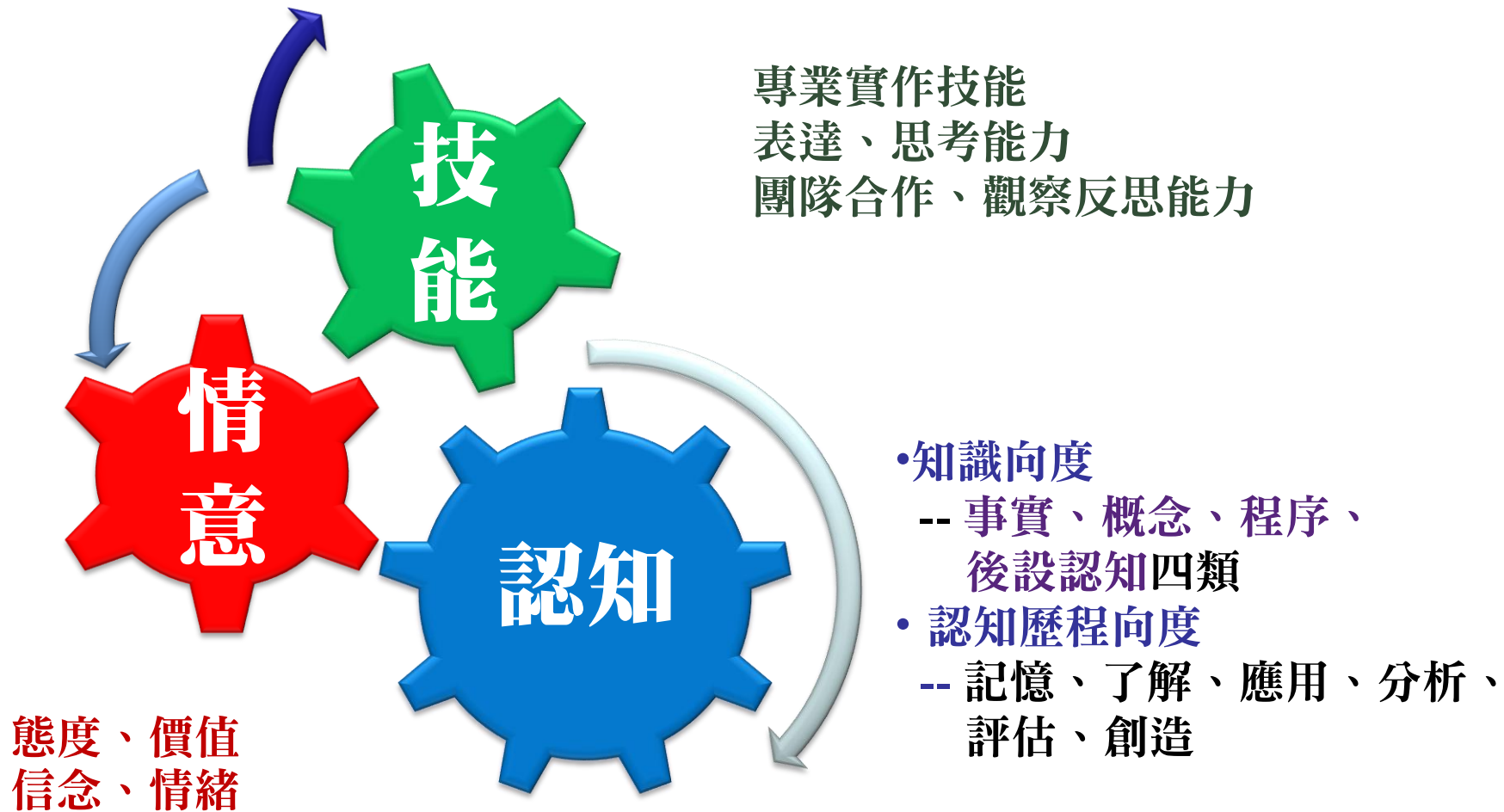


各項評分工作的實務挑戰

- 考試作業份數太多，**批改標準不一致**
- TA 與教師所認定的標準不一樣
- 學生不易得知沒有標準答案的批改準則
- 學生從成績中**無法知道哪些部分需要加強**
- 批改準則偏頗，缺乏全面性的評估
- 其他



學習目標面向



知識向度的類型

事實知識

- 術語的知識
- 特定細節和元素的知識

程序知識

- 特定學科的技能 and 演算的知識
- 特定學科技術與方法的知識
- 運用規準的知識

概念知識

- 分類和類別的知識
- 原理和通則的知識
- 理論/模式/結構的知識

後設認知知識

- 策略知識：知術(方法)
- 認知任務知識：知事
- 自我知識：知人



以學習表現為基準，清楚反應學生的學習成效。

明訂各項評量標準並配合評量學生的表現指標。

呈現方式：

評量標準 (criteria)：該學習行為所需具備的各個面向

表現指標 (performance)：定義各面向表現之優劣指標

Rubric 形式：

整體型：對學生整個學習表現進行整體性的評估

分析型：細分若干指標項目，各自訂定評量標準

整體型 Rubric: 書面專題報告

分數表現	詳細表現準則
A 非常優異	資料蒐集非常豐富 個人想法深入有創意 文字描述正確、清晰具邏輯性
B 表現普通	資料蒐集大致完整 呈現個人想法，但缺乏深入思考 文字描述清晰，些許錯別字
C 有待改進	資料蒐集略顯不足 毫無個人想法，欠缺思考分析 文字描述紊亂，錯別字甚多

分析型實例：生物統計學報告作業

項目	優良：A+~A-	普通：B+~B-	待改進：C+~C-
設立假說 × 1	題意描述詳細清楚 虛無假設正確詳盡 單位註明詳細	題意描述不甚清楚 虛無假設正確 單位註明無誤	題意描述交代不清 虛無假設有誤 單位註明錯誤
算式 × 3	選用分析方式正確 詳細交代分析方法	分析方式無誤 未詳述為何用此方法	選用分析方式錯誤 未交代為何用此法
計算推導 × 3	演算過程正確 算式推導過程詳細 位數選取合理 算式陳列清晰	演算過程無誤 算式推導不甚詳盡 位數選取合理 算式陳列不整齊	運算過程錯誤 算式推導含糊位數 選去過少 算式陳列凌亂
查表 × 1	選取正確查表 採用參數正確	選取正確查表 採用參數有誤	選取錯誤查表 採用參數有誤
結論 × 2	顯著性判斷無誤 結果陳述詳盡	顯著性判斷無誤 結果陳述不清楚	顯著性判斷錯誤 結果陳述不清處

Rubric 建立步驟

1. 參考Rubric範例. 選擇最適合課程性質與評量目的
2. 建立定義清楚的評估標準、子標題
3. 決定三或四個不同等級的標準
4. 訂出每個等級的分數範圍
5. 細項敘述的文字要明確、淺顯易懂
6. 尋求同儕的回饋，不斷修正
7. 在上課時與學生溝通、確認同學瞭解評估的標準

Rubric design	Levels of performance (scale)			
	4	3	2	1
Criterion 1	(Yes, and more!)	(Yes!)	(Yes, but...)	(No)
Criterion 2	Performance descriptors			
Criterion 3				
⋮				

[] = [] = []

Rubric 如何建立-1

1. 參考 Rubric 範例. 選擇最適合課程性質與評量目的

- Rubrics for Assessment
(www.uwstout.edu/soe/profdev/rubrics.cfm)
- Assessment Rubrics
(edtech.kennesaw.edu/intech/rubrics.htm)
- Rubric Maker Tools
(www.teach-nology.com/web_tools/rubrics/)
- The Incredible Art Department--Rubrics
<http://www.princetonol.com/groups/iad/links/toolbox/rubrics.html>
- Evaluation by Rubrics
(www.carla.umn.edu/assessment/vac/evaluation/p_1.html)

Rubric 如何建立-2



**設定學習目標
與評分方式**

Rubric 如何建立-3

建立定義清楚的評估標準、子標題

項目	題目關連性	前言方法	結果討論	簡報呈現	表現應答

Rubric 如何建立-4

決定三或四個不同等級的標準, 訂出每個等級的分數範圍

項目	題目關連性	前言方法	結果討論	簡報呈現	表現應答
優 A 10~9					
佳 B 8~7					
普通C 6~5					
待改進D 4~3					

Rubric 如何建立-5

細項敘述的文字要明確、淺顯易懂

項目	題目關連性	前言方法	結果討論	簡報呈現	表現應答
優 A 10~9			資料蒐集完整 分析精確 討論深入層次 分明.		
佳 B 8~7			資料蒐集與分 析尚可.討論 大致都有提到		
普通C 6~5			資料蒐集缺漏 分析稍有錯誤 討論深度不足		
待改進D 4~3			資料蒐集缺漏 分析錯誤.幾 乎沒有延伸討 論		

Rubric 如何建立-5

尋求同儕和學生的回饋，不斷修正

項目	題目關連性	前言方法	結果討論	簡報呈現	表現應答
優 A 10~9	題目與研究相 契合.研究與動 行有關. 題目吸引人	目的明確 假設合理 設計準確 具可行性	資料蒐集完整 .分析精確.討 論深入層次分 明.	文字圖表顯 示精準明確 大小配置佳 呈現效果佳	說明清楚流 暢.言簡意賅 時間分配佳 應答順暢
佳 B 8~7	題目與研究契 合.研究與動行 稍有落差	目的.假設大 致合理.設計 可以更完整 可行性適中	資料蒐集與分 析尚可.討論 大致都有提到	文字圖表顯 示正常.大小 配置適中.呈 現效果尚可	說明大致清 楚.時間分配 恰當 應答適中
普通C 6~5	題目有落差.研 究與動行關連 較少.引不起興 趣	目的不明 假設不合理 設計有瑕疵 可行性低	資料蒐集缺漏 .分析稍有錯 誤.討論深度 不足	文字圖表大 小顯示效果 差.呈現效果 普通	說明稍有不 明.時間分配 普通.應答不 完全
待改進D 4~3	題目與研究無 關，研究與動 物行為也無關	目的不明 毫無假設 設計瑕疵多 可行性差	資料蒐集缺漏 分析錯誤.幾 乎沒有延伸討 論	文意不明 圖表沒重點 大小配置差 造成反效果	說明含糊 廢話一堆 時間分配差 回答沒重點

整體型 Rubric 實例：口頭報告

5	優	學生清楚描述報告主題並提供足夠的理由說明它的重要性。有充份的資訊支持結論，演說方式吸引觀眾注意。學生全場保持與觀眾眼神接觸，可以明顯感受到對主題的準備與熱情。運用視覺輔助工具讓報告更具說服力。使用合宜的資訊清楚回答觀眾的問題。
4	佳	學生描述報告主題並提供足夠的理由說明它的重要性。有適量的資訊支持結論，內容大致正確。可以感受到對主題的準備與熱情。運用視覺輔助工具，清楚回答觀眾的問題。
3	好	學生描述報告主題與結論，但所提供的支持證據不夠周全。內容大致正確，可以約略感受到對主題的準備與組織。提及視覺輔助工具，但並未實際運用。試圖回答觀眾問題。
2	普	學生描述報告主題，但未能完整說明，且沒有提供結論。內容可以理解但有錯誤，無法感受到對主題的準備與組織。未提及或運用視覺輔助工具。僅以最基本的答案回答觀眾問題。
1	劣	學生未能描述報告問題及其重要性，主題不明確且沒有適宜的結論。內容難以理解，沒有準備且缺乏組織。僅以最基本的答案回答觀眾問題，或未能解答。

分析型 Rubric 實例：評量解決問題能力

	優	佳	可	劣
瞭解問題	瞭解研究問題，且能將其連結至真實世界的情況。	瞭解（能夠解釋）研究問題，並進行下一步。	需要旁人協助釐清，才能了解研究問題。	在開始之前，需要對於研究問題有完整的解釋。
蒐集資訊	蒐集許多資訊，且全部與主題相關。	蒐集一些基本資訊，大部分與主題相關。	蒐集非常少的資訊，其中有些與主題相關。	沒有蒐集任何與主題相關的資訊。
解決問題的方式	腦力激盪產生許多策略，並對每個策略有適當執行方法。	腦力激盪產生數個策略，並決定其中一個適當的策略。	腦力激盪產生少許策略，需藉由協助選擇適當的策略。	只針對問題設計策略，需要藉由協助評估策略。
施行計畫	挑戰自我，利用新的方法解決問題。	利用設計解決問題，並進行適當的修改。	雖然解決問題，但並未進行修改。	使用不適當的策略解決問題。
評估結果	建議其他的修正或運用方法於其他狀況，設計專屬的評估標準。	比較實際以及預期結果。	雖未透過協助，但對解決方法僅有有限的評估。解決方法僅適用於特定問題。	需要藉由協助評估結果。
溝通/呈現結果	總結解決方法，並描述這些方法可以如何被使用於其他狀況。	使用專業語言與符號（包括特定測量法與計量），解釋某一方法比較好的原因。	使用與研究問題相關的術語，解釋所發生的狀況。	利用簡單的方式解釋所發生的狀況。

分析型 Rubric 實例：評量創造研發能力

	進階	中級	初級
質疑並挑戰既有想法	能夠指出尚未被發現的問題。從不同角度檢視/挑戰既有想法。	獨立思考。挑戰傳統想法。	提出「為什麼/如何做」的問題。以不落巢臼的方式回應想法、問題、任務。
連結	產生原創的想法。說明具創意的洞見。以新穎的方式傳達新想法。	重新詮釋所學或應用所學於新的情境中。將通常沒有關係的事情互相連結。	只看到表淺明顯的連結。
發展想法	投入於具突破性的想法。創造新的想法、概念與可能性。創造新的架構以組織想法。	從多元角度預見問題與挑戰。醞釀新想法與概念。對於過程加以改進。挑戰現有想法的極限。	利用他人想法或結論。在現有的架構中思考。用不一樣的方式解釋新想法、概念。
將概念轉換為行動 (如作品或實踐)	投入於持續的改進。藉由尋求外在評估，有目的地求證新的想法。	整合想法與概念。採用工具與資源，促進發展中的概念。架構與測試模型或原型	利用技能與知識。利用資源以將概念轉換為行動。投入於設計過程。

- **Performance Criteria:** *What is being evaluated?*
 - Are the performance criteria linked to standards?
 - Are there a manageable number of performance criteria? 3-5?
 - Are the performance criteria measurable/teachable?
 - Do performance criteria match the objectives?
 - Are the performance criteria clearly stated with a meaningful label?

- **Scales and Levels of Proficiency:** *Degrees of quality?*
 - Is there an even number of levels, i.e., 4 to avoid middle scoring?
 - Does the highest level represent exemplary performance?
 - Is each level on the scale meaningful and non-judgmental?

- **Descriptions:** *What would success look like?*
 - Are they written in student language? Clear and understandable?
 - Are they positively stated?
 - Are the differences in descriptions observable? Clearly stated expectations?
 - Is there a progression of differences among the descriptions?

Rubric 總體檢視

- Does the rubric have these elements: **Performance Criteria**, **Scale**, **Levels of Proficiency**, and **Descriptions**?
- Is the rubric **manageable** and **practical** to use by students and teachers?
- Have others (e.g., peers and students) reviewed this rubric?

使用評量指標的好處

對教師而言：

- (1) 客觀評量，避免主觀的成見或既定印象。
- (2) 能夠明確回應學生對於評分高低的疑問。
- (3) 能夠節省教師在評估作業及提供回饋所需時間。
- (4) 能夠幫助教師檢視並調整課程目標與教學行為。



(2) 定義各級表現指標
Performance



(1) 明訂幾個評量標準
Criteria



對學生而言：

- (1) 知道成績的評量標準及各等級範圍，引導正確有效學習。
- (2) 可藉此知道自己的優缺點，及未來應該朝何種方向改進。
- (3) 可發展自我評估的能力，並為自己的學習負責。

總結

- 好的評量方式是有效、可靠且切合學習目標的
- 隨時間成長的評量，教學及學生學習相結合
- 使用敘述評語、豐富學習的整合
- 評分為教學提供情報、激發學習潛能
- 鼓勵學生成為主動的自我評量者

謝謝參與！



以教育彩繪臺灣的未來