

第3章 多元化評量 20-41

教學與評量是現代教育系統中的兩大部分，兩者之間有互動，但其關係仍是教學為主，評量為輔。評量為教學的一項後續 (follow-up) 活動，用以了解學生學習成果，及調整或擬定日後教學策略。但由於學生、家長、甚至許多老師未能了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、在全班的位序或排名等；再加上傳統評量方式設計不周，不夠多元化，未能涵蓋教學內容的各個層面，形成教學與評量之間的落差，使得評量的角色有機會易客為主，對教學產生反向的導引作用 (backwash)，亦即所謂的「考試領導教學」。所幸，這種反向的導引作用，雖然可能為阻力，亦可能成為助力 (Hughes, 1989)，端看如何因應。以國內大學聯考為例，將英文作文納入考試，進而帶動國內英語老師及學生對英文作文的重視，便是因勢利導，使一個反向的導引作用成為助力的最佳例子。基於多元評量的重要性，本章將釐清一些多元評量的基本理念，並探討如何將其落實在九年一貫新課程的英語科中。

3.1 評量的定義：評鑑、評量與測驗 20

許多外語教學的學者如 Valette (1994)，對 evaluation (評鑑/評量)、assessment (評量)、testing (測驗) 並不作特別區分；更多的學者僅作二分法，一類是層次較低的 testing，另一類則層次較高，有的學者用 assessment 一詞，如 Bailey (1998)，有的用 evaluation，如 Genesee and Upshur (1996)。

事實上，這三者之目的及範圍上，皆有所不同 (Brown, 1989; Lynch, 1996)。testing 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、隨堂考等皆屬此類。assessment 範圍較廣，以測驗、面談 (interview)、實作或工作計畫 (project)、檔案評量 (portfolio assessment) 等方式，診斷學生的語文能力。evaluation 層次更高，係“有系統地蒐集相關資料以改善課程、評估其成效、了解學生的態度” (Brown, 1989)，甚至可包括“評估測驗試題的品質” (Alderson, Clapham, and Wall, 1995)。在中文裡，此三詞可分別以『測驗』『評量』『評鑑』翻譯，界定明確，不至於有所混淆。不過，如上所述，有一些常用詞語如 formative evaluation 及 summative evaluation，因許多外國學者沿用已久，故其本質雖為評量，英文部份雖然有時仍沿用 evaluation，中文則以「評量」稱之。