

壹、前言

課室評量 (classroom assessment) 維繫著教學品質之良窳。Black 和 Wiliam (2010) 回顧 250 篇有關課室評量的文獻，發現教師採用良好的評量實務，能有效促進學生的學習理解，其效果值 (effect size) 介於 .4 到 .7 之間，比起大多數的教學介入更能有效提升學生的學習成就。而一個好的課室評量，應充分連結學習目標與教學方法，以支持教師的教學和學生的學習 (Taylor & Nolen, 2008)。就整個教學歷程來看，從起點行為的診斷、教學目標的設定、到教學方法的採用，皆與課室評量息息相關。教師透過良好的評量方式，不僅可以瞭解學生已學會哪些知識技能 (assessment of learning)，進而可在學生學習的可能發展區 (the Zone of Proximal Development, ZPD) (Vygotsky, 1930-1935/1978)，選擇對學生有意義且能理解的學習內容，經由適當的教學引導與學習回饋，以促進學生的學習 (assessment for learning)，達到預定的學習目標。

就評量目的而言，一般可將評量分為著重在評定學生學習成果的總結性評量 (summative assessment)，以及著重學習診斷與教學改進功能的形成性評量 (formative assessment)。總結性評量雖具有評鑑教學計畫的實施成效、評量學生的學習成就、區分學生等第與篩選人才等功能，但卻無法提供足夠的訊息，作為學生精進學習與教師改進教學之參考。反之，形成性評量是適用於教與學過程中的評量，可即時確認學生的學習狀況，進而透過適當的協助，以促進學生的學習 (Bell & Cowie, 2001; Black, Harrison, Lee, Marshall, & Wiliam, 2004; Black & Wiliam, 1998)。在課堂中，若能將評量融入於教學過程，讓每個教學活動都伴隨著評量任務，不僅學生可透過評量任務的練習，促進對科學概念的理解 (Treagust, Jacobwitz, Gallagher, & Parker, 2001)，教師也可透過學生的反應，瞭解學生的學習情形，以決定何時教導學生新的概念與決定採用何種教學活動，以維持學生的學習動機和興趣 (Bell & Cowie, 2001)。

基本上，在整合評量於教學的過程中，教師對學生的評量方式、對評量結果的詮釋，以及對學生的回饋品質，皆攸關學生學習的成效 (Bell, 2007; Black et al., 2004)。可惜的是，在教師專業養成的階段，有關評量的相關課程並不受重視，導致教師的評量素養普遍不足 (MacLellan, 2004; Mertler, 2004)。一項於美國全國性的教師評量素養抽樣調查發現，在所有評量素養的向度中，教師在「與學生、家長或其他教育者溝通評量結果」的能力最低 (Plake & Impara, 1996)。換句話說，教師對於評量結果的解釋是有困難的，而這也直接影響到評量結果在後續的教學運用。

為了提升教師的評量實務能力，許多研究也紛紛著手開發教師評量素養培育課程。例如：Mertler (2009) 根據「學生教育評量的教師能力標準」(the standards for teacher competence in the educational assessment of students) 設計一套為期 2 週的在職教師課室評量工作坊，經實驗研究

發現，該工作坊課程能增進教師的評量知識與技能。Sato、Wei 和 Darling-Hammond (2008) 的 3 年縱貫研究也發現，教師透過參與全美專業教學標準委員會 (the national board for professional teaching standards) 的認證過程，可提升其評量實務的能力，包括採用各種評量的能力和運用評量訊息來支持學生的學習等。在國內，王子華、范雅晴和王國華 (2008) 運用「個人化」和「情境化」的網路評量和試題分析系統 (web-based assessment and test analysis) (Wang, Wang, Wang, Huang, & Chen, 2004)，發展的數理科在職教師評量素養培育模式，亦有助於提升教師的評量知能。而此模式運用於培育職前生物科教師的評量素養，也得到類似的成效 (Wang, Wang, & Huang, 2008)。

然而，目前的教師評量素養培育課程多著重在提升教師的評量知識，有關於如何在課堂中運用評量的回饋，以促進學生學習的技術，則較少有相關的文獻。就課堂中的形成性評量而言，評量結果的回饋乃扮演著促進學習的關鍵性角色 (Black et al., 2004)。良好的評量回饋不僅能讓學生知道自己的優缺點 (Black & Wiliam, 1998)，也可讓學生知道自己的表現是否達到預定的學習目標 (標準參照)、自己的表現和其他同學的比較 (常模參照)，以及自己的學習成長情形 (自我參照) (Bell, 2007)。不過，並非所有的回饋都可以促進學生的學習 (Hattie & Timperley, 2007; Kluger & DeNisi, 1996)。Kluger 和 DeNisi (1996) 的後設分析研究發現，雖然多數的回饋都有助於學習，且平均效果值達 .41，但仍有三分之一的回饋會降低學習表現。他們發現，當回饋只是單純的獎勵或評判，將降低回饋對學習的助益。反之，有助於學習的回饋方式，則多是將回饋聚焦於告訴學生如何達到預定目標的改進方法 (Lipnevich & Smith, 2009)。Hattie 和 Timperley (2007) 回顧了探討回饋效果的文獻，歸納出有效的回饋必須要能回答學生三個問題：一、學習的目標為何？二、要到達目標的途徑為何？三、目前應採取何種行動才能更接近目標？以 Vygotsky (1930-1935/1978) 的 ZPD 概念來說，良好的評量應該要能辨識出學生學習的 ZPD 範圍，亦即指出哪些知識概念是學生既有能力所應該學會的，哪些知識概念是學生在教師或同儕協助下，可以學會的，而評量回饋的目的就像是為學生搭建學習的鷹架，透過教師或同儕的協助，縮小學生目前已學會和未來預期學會的目標之差距，支持學生一步一步地朝學習目標邁進 (Shepard, 2005; Shute, 2008)。

可惜的是，在升學考試制度的影響下，教師很少運用課室評量的回饋技巧來幫助學生的學習。研究發現，教師為了配合國家考試的內容與形式，會傾向採用總結性成就測驗，著重於評量學生是否學會國家的課程目標，而較少運用形成性評量，提供學生學習改進方向 (Gioka, 2009; Ogan-Bekiroglu, 2009)。此外，教師因擔心自己的評量素養和學科專業知識不足 (Wang, Kao, & Lin, 2010)，以及編製測驗與實施評量的時間有限 (Graham, 2005)，往往採用傳統的教師自編成就測驗進行課室評量。以當前的臺灣教育現場為例，教師在課堂中所進行的評量，仍是習慣在單元教學後，依據課程綱要所訂定的能力指標和單元主題內容，自行編製評量測驗進行考試，並在考試結束後，採用測驗總分評定每位學生的成績，再逐題訂正