

壹、緒論

歷經多年的倡議與規劃，十二年國民基本教育（以下簡稱十二年國教）自 2011 年起正式啟動，我國的教育工程正式邁向一個新的里程碑。為了順利推動此重大教育政策，主管機關提出七大工作要項與十項方案，在眾多措施中又以入學制度及評量方式的變革最引人關注。依據高中、高職及五專免試入學實施方案，高中、高職及五專多元入學管道將於 103 學年度起整合為「免試入學」與「特色招生」兩種管道，免試入學意指學生不需參加入學考試，經錄取報到後即可進入高中、高職或五專就讀（教育部，2013a），特色招生係指具學科或術科性向的學生參加入學測驗，學校據此招收與其發展特色相符之學生（教育部，2013b）。自 103 學年度起免試入學名額比例將逐年提高且考試入學名額逐漸降低（分別達 75% 以上及 25% 以下），由此可知，未來大部分國中畢業生都將透過免試管道升學。

十二年國教歷經多年規劃，但在眾所矚目的入學與考試制度的變革，卻始終未能獲得共識。主張降低升學壓力的團體與主張維持競爭力的團體，各有論述，各持立場，讓教育政策在天平的兩端擺盪，難以定奪。目前定調以免試為主的升學制度固然有助於緩解升學壓力，卻也不免產生學生學習動機下滑、學習成就品質低落之隱憂。如何制定出能同時平衡降低考試壓力，但又能維持適當競爭力的入學與考試制度，為十二年國教的重要挑戰之一。除了學校以外的外部考試，如何發揮學校內部班級評量，透過低利害（low-stakes）考試評量以帶領學校教學方式的變革，落實培養學生高層次與多元學習能力，以及活化教師教學的目標，是十二年國教的另外一個迫切面臨之課題。

實施以標準為本位之評量方式（standards-based assessment, SBA，以下稱標準本位評量）為許多先進國家與地區（如美國、澳洲、英國與香港等）近年來的教育評量改革趨勢，這些國家（或地區）一方面將標準本位評量應用在外部考試以監控學力，另一方面也將其應用在班級評量，透過促進課程標準、教學與評量三者的對應，提升教師評量專業與學生學習成效（Lyon, 2011; Martone & Sireci, 2009）。以美國為例，2001 年公布的《沒有孩子落後法案》（No Child Left Behind Act, NCLB）要求各州政府確保學生每年都有適度的學習進展，並以每位學生在核心學科上皆能達到優秀（proficient）等級為目標（Linn, Baker, & Betebenner, 2002），為了達到《NCLB 法案》的要求，許多州政府開始訂定適用該州的評量標準且每年施測大型標準化測驗（large-scale standardized assessment）以監測學生整體成就水準。在班級評量部分，美國加州（California State）針對英語能力制定了該州的評量標準，並於部分公立學校實施標準本位班級評量（standards-based classroom assessment, SBCA），希望藉此提升教師教學與評量專業並回饋到學生學習成果（Llosa, 2011）。

本研究之目的在於闡述標準本位評量的理念，以及如何透過標準本位評量促進十二年國教之入學與考試制度變革，進而提升實現十二年國教理想的可行性。

貳、標準本位評量

一、標準本位評量之意義

標準本位評量是指在評量前制定系統化的評量標準並據此發展評量工具，評量後將結果對照到評量標準以將學生學習表現劃分成不同表現等級，進而解釋與說明該生學習進展之評量歷程。建置評量標準是實施標準本位評量的先決條件，為使評量結果能用來瞭解學生是否達成教師的教學目標，評量工具所測量的知識概念應明確對應到教師的授課內容。在實務上，教師應根據官方制定的課程規範（例如我國的課綱）設立教學目標、設計課程與教材、進行教學活動等，因此，若能同樣根據課綱與能力指標建構評量標準，將有助於串連評量與教學，評量結果更有助於瞭解學生是否具備課綱期望的學習成就（宋曜廷，2012a）。

二、標準本位評量之組成要素

評量標準由內容標準（content standards）與表現標準（performance standards）所構成，內容標準意指希望學生具備的知識與展現的技能（Hambleton, 2001），例如，我國數學科課綱針對七年級學生，期望其能具備認識數與數線、理解因數與倍數或一元一次方程式等知識。表現標準則是說明在內容標準所條列的學習內容中，學生能展現相關的知識與技能達到何種程度（Linn & Herman, 1997）。由於不同學生在教學活動後可能會呈現出不同的表現程度，因此須設立適當數量的表現等級（performance levels），用以大致區分學生學習進展的差別，例如，美國「國家教育進展測驗」（National Assessment of Educational Progress, NAEP）將學生學習表現區分為進階（advanced）、優秀（proficient）、基礎（basic）與未達基礎（below basic）等，共四個等級。此外，需針對各個等級撰寫表現等級描述（performance level descriptors, PLDs），讓教師可以進一步瞭解不同等級學生的典型表現或最低門檻水準。

透過將學生表現對應至預先建置的評量標準，教師便可界定出學生在特定學習內容上的表現等級，此種參照到標準以產生成績等級的評量方式，改善了常模參照（norm-referencing）的部分缺失。在常模參照中，透過與群體中其他人的測驗分數進行比較，可以獲得該生在群體中的百分等級（percentile rank, PR）以瞭解相對地位表現，該量化分數雖有利於分發入學，但也被認為助長了對測驗分數分分計較的現象。此外，量化分數未能提供學生學習成果的訊息，教師難以根據評量結果進行診斷與補救，例如，我們可以知道 PR 值為 40 的學生其表現低於群體中六成的學生，但無法指出該生缺乏何種知識或哪些技能需要補強。相對地，在標準本位評量中，對該生表現等級的界定不需要與他人比較，評量標準可提供各表現等級的質性表現描述，因而能指出該生目前具備的知識與技能。因此，以標準為本位的評量方式不但可降低競爭壓力、讓學生專注在提升自身學習成果，更能拓展評量結果的解釋與應用，提供