

教育資料與研究雙月刊

第68期 2006年2月 頁79-100

基本能力評量跨國發展經驗之比較

張鈿富 王世英 吳慧子 周文菁

摘要

本研究以國際間發展的基本能力評量趨勢為探討對象，參考主要國家的經驗與特點，以收他山之石。基於此，本研究主要目的包括：

1. 比較各主要國家學生基本能力評量的內涵。
2. 探討各主要國家推動學生基本能力評量的策略。
3. 分析各主要國家學生基本能力評量的呈現方式。
4. 分析我國推動此類學習評量的配套措施。

本研究「主要國家」界定為美、英、OECD等先進國家或組織為主，以他們發展的學生評量為比較基礎，分別探討各評量計畫的優缺點，以做為我國發展類似評量系統的參考。

關鍵詞：基本能力評量、NAEP、英國國家課程、OECD/PISA

張鈿富，國立暨南大學人文學院院長

王世英，國立教育資料館館長

吳慧子，國立暨南大學教育政策與行政學系助理教授

周文菁，國立暨南大學教育政策與行政學系助教

電子郵件為：dhchang@ncnu.edu.tw；wang@mail.nioerar.edu.tw；julinchou@yahoo.com.tw

來稿日期：2006年1月12日；修訂日期：2006年2月9日；採用日期：2006年2月21日

Comparing the Experiences of Developing Basic Student Performance Assessment in Different Countries

Dian Fu Chang Shih Ying Wang
Hui Tzu Wu Wen Ching Chou

Abstract

This paper focuses on the trend of the basic performance assessments developed in different countries. Reviewing the experiences and traits in different assessment systems, the authors try to summarize a proposed framework for further development in Taiwan. Based on the ideas, the main purposes are as follows °

- (a) Comparing the contents of different basic performance assessment frameworks in different countries.
- (b) Discussing the different strategies for implementing the student basic performance assessments.
- (c) Analyzing the report styles of different assessment systems in different countries.
- (d) Analyzing the strategies for developing the student basic performance assessment system in Taiwan.

In the paper, the main country or organization defined as United States of American, British, and OECD. Based on their student basic performance assessments, the authors addressed the different traits for further developing the basic student performance assessment in Taiwan.

Keywords: basic performance assessment, NAEP, National Curriculum of British, OECD/PISA

Dian Fu Chang, Prexy, College of Humanities, National Chi Nan University

Shih Ying Wang, Director, National Institute of Educational Resources and Research

Hui Tzu Wu, Assistant Professor, Department of Educational Policy and Administration

Wen Ching Chou, Assistant, Department of Educational Policy and Administration

E-mail: dhchang@ncnu.edu.tw ; wang@mail.nioerar.edu.tw ; julinchou@yahoo.com.tw

Manuscript received: January, 12, 2006 ; Modified: February, 9, 2006 ; Accepted: February, 21, 2006

壹、緒論

學習品質的確保是近年來各先進國家教育發展的重點，爲了確保學校教育的品質，各國紛紛進行了「基本能力」或是「基本素養」的評量。

爲因應國際教育發展的趨勢，未來我國學生基本能力評量應何去何從？因此，進行跨國性學生基本能力評量之比較研究有其必要性。如何展開系統性研究、掌握最新的發展趨勢，以做爲我國發展類似評量的參考，爲本研究發展的重點。

一、研究緣起

爲了教育品質，各地區或是各國都有因應的策略。例如，美國評量學生成就的代表爲「國家教育發展評量」(NAEP) 所發展，是由教育測驗服務社(ETS)所發展的聯邦補助計畫。NAEP採用大量的、有代表性的學生樣本去建立學生學習成就的趨勢。此一評量的結構和目標經由課程專家、教師、家長、商業領導者和政策制定者的共識發展而成，反應了學生應該知道和可以做的廣泛能力。成就評量的焦點放在九、十三、十七歲的學生。

英國亦把學習評量納入國家課程發展的一環。這些學校外部的評量設計都以建立標準(standard)或是標竿(benchmark)爲主，以確保學校教育系統品質。

歐盟推動教育「品質保證」(quality assurance) 機制，其中教育結果的部分以「經濟與合作發展組織」(OECD)所發展的PISA (Program for International Student Assessment) 評量結果爲主，來瞭解各參與國家十五歲學生的學習表現。

二、研究目的

由於學生素質確保的重要性，而學習評量是維繫教育品質的一個重要手段。尤其是學生基本能力的評量，與未來公民的素質、從事經濟發展所需的人力素質息息相關。在此一前提下，本研究以國際發展這類評量的趨勢爲主，參考各國的經驗與特點，以收他山之石之效。基於此前提，本研究探討之目的如下：

- (一) 比較各主要國家學生基本能力評量的內涵。
- (二) 探討各主要國家推動學生基本能力評量的策略。
- (三) 分析各主要國家學生基本能力評量的呈現方式。
- (四) 分析我國推動此類學習評量的配套措施。

本研究的「各主要國家」界定爲美、英、OECD等先進國家或組織爲主，以他們發展的學生評量爲比較的基礎。

三、研究待答問題

根據以上研究目的，本研究主要的待答問題如下：

- (一) 各主要國家學生基本能力評量的內涵與優缺點爲何？
- (二) 各主要國家推動學生基本能力評量的策略爲何？
- (三) 各主要國家學生基本能力評量的呈現方式有何借鑑之處？
- (四) 我國推動學生基本能力評量的可能配套措施爲何？

四、研究設計

本研究透過文獻分析，以比較研究為主。首先，蒐集各主要國家有關學生基本能力評量的文獻，「各主要國家」以英、美、OECD等主要國家或組織為主，就目前已發展的學生基本能力評量為文獻蒐集的範圍。其次，就蒐集的文獻進行比較，比較各主要國家學生基本能力評量的內涵，並分析這些內涵的代表意義，以做為我國發展類似能力評量的參考。初步研究成果提出後並進行專家的諮詢，以確保研究結果的正確性。最後，提出發展我國學生基本能力評量的配套措施，以做為教育部相關政策訂定之參考。具體而言，本研究的執行流程如圖1：

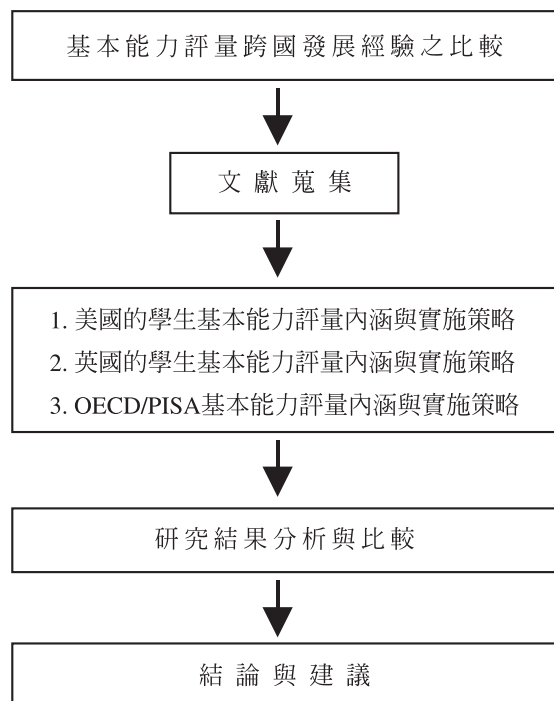


圖1 研究流程

貳、美國學生基本能力評量

「國家教育發展評量」The National Assessment of Educational Progress (NAEP，以下簡稱NAEP)是美國評量學生成就的代表，本節介紹此一評量的內涵，並以NAEP為例，分別說明四年級、八年級以及十二年級實施的閱讀能力評量(reading)、數學能力評量(mathematics)與科學能力評量(science)。

依據美國法律規範 (U.S. Department of Education, 2005)，全國教育統計中心 (National Center for Education Statistics，以下簡稱NCES) 的最高行政長官應負責執行NAEP這個政策。自1988年以來，NAEP在全國評量管理委員會 (National Assessment Governing Board，以下簡稱NAGB) 所制定的政策指導之下，平穩的執行其功能。NAGB是由NCES所指定組成的一個管理委員會，成員包括教育專家學者、推選出的官方代表、家長、以及一般民衆代表。平時美國民衆所聽見熟悉的名詞The Nation's Report Card，其實即是NAEP的一種親切用語。The Nation's Report Card採用一種非技術性(即易懂的用詞)方法來介紹NAEP主要內容。The Nation's Report Card的目的是為了老師們、家長們、和一般社會大眾，瞭解關於全國評量的重要資訊。

一、組織

當談到與NAEP相關的組織時，NCES與NAGB是兩個關鍵的單位。全國教育統計中心 (NCES) 是美國教育部 (U.S. Department of Education) 直屬教育科學部門 (Institute of Education Sciences，以下

簡稱IES) 的四個所屬研究單位之一。NCES 負責執行此項由上而下的計畫 (mandated project)：「國家教育發展評量」(NAEP)。

全國評量管理委員會 (NAGB) 於 1988 年經國會通過，是一個獨立於美國教育部的機構，目的在於制定及領導 NAEP 之相關政策 (The Nation's Report Card, 2005a)。NAGB 的成員包含聯邦的 (federal)、州政府的 (state)、以及地方的官員 (local officials)，此外還有教育學者專家、家長、商業領導者、與社會一般民衆的成員代表。

二、目的

NAEP 由實施評量的地方範圍可簡單分成：全國的 (National NAEP)、各州的 (State NAEP)、以及地方的 (NAEP Trial Urban District Assessment) 評量 (The Nation's Report Card, 2005a)。NAEP 主要目的是：

- (一) 反映美國學生在主要課程領域上應該知道和可以做的廣泛能力 (to discover what American students know and can do in key subject areas)；
- (二) 測量長時間範圍內的教育發展情形 (to measure educational progress over long periods of time)。

三、類型

爲了達成這些目的，NAEP 的計畫包括兩種重要評量類型 (The Nation's Report Card, 2005a)：(1) 主要評量 (Main NAEP) 與 (2) 長期發展趨勢評量 (Long-term Trend NAEP)。

主要評量 (Main NAEP) 注重學生現

在學習的內容是什麼 (what students are learning today)。一般 NAEP 所呈現的教育評量結果大多屬於主要評量 (Main NAEP)。Main NAEP 評量的內容每十年都會重新檢視與更新。每兩年，全國的四年級、八年級與十二年級學生需接受閱讀能力 (reading) 與數學能力 (mathematics) 評量；同樣的，每兩年，各州的四年級與八年級學生需接受閱讀能力與數學能力評量。每四年，全國的四年級、八年級與十二年級學生需接受科學 (science) 與寫作 (writing) 評量；同樣的，每四年，各州的四年級與八年級學生需接受科學能力與寫作能力評量。

長期發展趨勢評量 (Long-term Trend NAEP) 主要爲全國的施測，屬於全國性評量。主要評量閱讀 (reading) 能力與數學 (mathematics) 能力的長期發展趨勢。每四年，NAEP 針對九、十三和十七歲的學生實施數學 (mathematics) 與閱讀 (reading) 評量，採用的評量內容與架構是修正自 1969 年以來的評量測驗題目。因此，長期發展趨勢評量 (Long-term Trend NAEP) 的每次結果可以拿來比較，如表 1。

除了 Main NAEP 與 Long-term Trend NAEP 兩種主要評量類型之外，還有一些特別議題的 NAEP 評量研究，例如目前正在進行中的高中學習成績研究 (High School Transcript Study) 及科學基礎的評量 (Technology-Based Assessment) 等 (The Nation's Report Card, 2005a)。

NAEP 實施評量的課程領域有十一個 (administers assessments in 11 subject areas)：藝術 (The Arts)、公民 (Civics)、經濟 (Economics)、外國語文

(Foreign Language)、地理 (Geography)、數學 (Mathematics)、閱讀 (Reading)、科學 (Science)、美國歷史 (U.S. History)、世界歷史 (World History)、以及寫作 (Writing)。較常見的評量領域為其中的八個能力：閱讀、數學、科學、寫作、美國歷史、公民、地理、和藝術。

表1 長期發展趨勢評量 (Long-term Trend NAEP) 與主要評量 (Main NAEP) 的比較

	長期發展趨勢評量 Long-Term Trend Assessment	主要評量 Main NAEP Assessment
目的 Purpose	每四年評量學生在數學與閱讀上面的表現，2004年公布目前最新的評量，2008年將在公布下一次的評量結果。	每兩年評量學生在數學與閱讀上面的表現，2005年公布目前最新的評量，其他的科目領域學生也會接受評量。
評量的內容 Content Assessed	評量的基本內容從第一次的評量（1971年第一次閱讀評量，1973年第一次數學評量）開始就維持保留不變，雖然自2004年有一些修正。	評量每十年會修正改變，來反應在全國學校課程上的調整改變。新的評量架構可以因應這些修正。
學生的抽樣 Students Sampled	全國性代表即選取九歲、十三歲與十七歲的學生。樣本人數比主要評量的樣本人數小。2004年的結果，僅限於公立學校的白人學生、非裔學生與西班牙裔學生。	選取四年級、八年級與十二年級的學生。偶數年評量全國性代表的學生，奇數年評量各州及城市地區的學生。抽取較大量的樣本來代表少數族群學生，例如亞裔及太平洋學生、美國印地安學生。
評量實施 Administration	每四年實施評量。	每兩年實施數學與閱讀能力評量。
評量結果呈現方式 Results Reported	結果呈現學生成就表現與長期的趨勢發展。成就表現的呈現方式利用評量分數，但並非使用學生成就的好壞來呈現結果。結果只報告全國性的資料。	評量結果的呈現方式為評量分數，並且呈現成就表現結果為基礎、熟練或是精熟。自1990年以來，結果的呈現包含全國性的、參與的各州以及其他的參與行政地區。

資料來源：The Nation's Report Card (2005b) .

四、實施評量時間

NAEP (The Nation's Report Card, 2005a) 並不呈現學生個人或是個別學校的評量結果，NAEP呈現的結果是特定領域的學習成就 (NAEP does not provide scores for individual students or schools; instead, it offers results regarding subject-matter achievement)。NAEP的評量是由一群經抽樣選出的學生來接受評量所代表的結果。實施NAEP可區分為全國性的評量 (National NAEP)，州層級的評量 (State NAEP)，以及都市地區的評量試驗 (NAEP Trial Urban District Assessment)。

全國性的評量 (National NAEP) 呈現的評量結果是代表全國性的評量結果，包括公立學校與非公立的學校，評量學生於四年級、八年級與十二年級的成就。自1990年起，加入州層級的評量 (State NAEP)，呈現各州的評量結果，目前評量學生以四年級與八年級的學習成就為主。都市地區的評量試驗 (NAEP Trial Urban District Assessment) 主要是因應No Child Left Behind Act，針對都市地區 (urban districts) 的學生實施NAEP評量。NAEP預計2005年到2017年的評量計畫時程表，如表2。

表 2 從2005年到2017年NAEP預計實施評量順序時間表
Schedule for the State and National Assessment of Educational Progress (NAEP) from 2005—2017

年	主要評量Main NAEP		長期發展趨勢 (Long-term Trend)
	全國的National	各州的State	
2005	閱讀 (reading) 數學 (mathematics) 科學 (science) 高中成績研究 (high school transcript study)	閱讀 (reading) (4, 8) 數學 (mathematics) (4, 8) 科學 (science) (4, 8)	
2006	美國歷史 (U.S. history) 公民 (civics) 經濟 (economics) (12)		
2007	閱讀 (reading) (4, 8) 數學 (mathematics) (4, 8) 寫作 (writing) (8, 12)	閱讀 (reading) (4, 8) 數學 (mathematics) (4, 8) 寫作 (writing) (8)	
2008	藝術 (arts) (8)		閱讀 (reading) 數學 (mathematics)
2009	閱讀 (reading) 數學 (mathematics) 科學 (science) 高中成績研究 (high school transcript study)	閱讀 (reading) (4, 8) 數學 (mathematics) (4, 8) 科學 (science) (4, 8)	
2010	美國歷史 (U.S. history) 公民 (civics) 地理 (geography)		

表 2 從2005年到2017年NAEP預計實施評量順序時間表
Schedule for the State and National Assessment of Educational Progress (NAEP) from 2005—2017 (續)

2011	閱讀 (reading) (4, 8) 數學 (mathematics) (4, 8) 寫作 (writing)	閱讀 (reading) (4, 8) 數學 (mathematics) (4, 8) 寫作 (writing) (4, 8)	
2012	世界歷史 (world history) (12) 外國語文 (foreign language) (12) Probe: (technological literacy) (special study)		閱讀 (reading) 數學 (mathematics)
2013	閱讀 (reading) 數學 (mathematics) 科學 (science) 高中成績研究 (high school transcript study)	閱讀 (reading) (4, 8) 數學 (mathematics) (4, 8) 科學 (science) (4, 8)	
2014	美國歷史 (U.S. history) 公民 (civics) 地理 (geography)		閱讀 (reading) 數學 (mathematics)
2015	閱讀 (reading) (4, 8) 數學 (mathematics) (4, 8) 寫作 (writing)	閱讀 (reading) (4, 8) 數學 (mathematics) (4, 8) 寫作 (writing) (4, 8)	
2016	藝術 arts (8)		閱讀 (reading) 數學 (mathematics)
2017	閱讀 (reading) 數學 (mathematics) 科學 (science) 高中成績研究 (high school transcript study)	閱讀 (reading) (4, 8) 數學 (mathematics) (4, 8) 科學 (science) (4, 8)	

資料來源：The Nation's Report Card (2005c)。

參、英國學生基本能力評量

英國在1988年七月通過教育改革法案，確立了國定課程及評量測驗的實施。由於近年來英國在世界舞台上倍感國際競爭的壓力，不但國力無法與日本、美國相提並論，歐洲共同市場的地位也受到德、法兩國的威脅。因此國定課程的實施是英國教育史上相當重要的里程碑，具有提升國力、重振往日聲威的重責大任。本節就英國國定課程設立目的、課程目標與科目及國定課程與評量發展等內涵做一探討。

一、國定課程設立目的

英國的國定課程設立基於下列的目的：

(一) 減少各校之間的歧異

在國定課程實施之前，英國的義務教育在各校之間存有很大的歧異。比如在一個地區，某個學校可能有完整的科學課程，另一個學校可能完全沒有。即使兩校均有的科學課程，也不能保證都以同樣的方式教學，甲校可能非常重視基礎科學，而乙校則偏重物理及地球科學，其他科目亦是如此。國定課程的實施可以確保所有學生學習廣博而均衡的學科內容。

（二）排除偏頗選擇

在英國中學裡常可看到偏頗的課程選擇，如女生通常會選生物而放棄物理，男生則多選科技而非家政，不少男生也常放棄外國語或音樂。國定課程的實施將使學生接受同樣良好而適切的課程內容。

（三）提升學習品質

評量的實施將有效監控學生在各階段的表現與各項目標上的進步，沒有通過的學生要加強補救，以保障學生獲得最起碼的知能。同時國定課程所訂的標準將引導年輕人有較高的預期及自信，成為較有能力的人。

（四）建立績效責任

教育是公共政策事務的一環，所以依據社會大眾期望所設計的共同課程，將使家長得以判斷孩子的進步情形及學校績效，地方教育當局更能評估其轄區學校的得失，甚至全國性評鑑亦得以進行，如此教育投資將不至浪費，績效責任得以建立。

（五）導引課程的聯繫

國定課程設立目的之一是改善小學和中學之間的連繫。在此之前，由於各校因地方當局的教育政策或各校的課程安排互異，教師面對來自不同學校的學生，常抱怨無法掌握課程內容，特別是一些有次序性的科目（如數學、科學、音樂等）更是困難。其次，小學的課程設計多不以學科為單位，但中學面對十六歲證書考試，必須以學科為主，且選修課程增多，課程難以連貫。國定課程的實施將可使義務教育的課程結構形成一個連續而一貫的整體，便於銜接。

（六）改善評鑑標準

國定課程的評量結合教師評量與標準

評量，兼具了「形成性」、「診斷性」、「綜合性」、「評鑑性」四種方式，是一種比較完整且符合人性化的設計。

二、課程目標與科目

（一）國定課程的目標

使個體具備成人生活所需的知識及技能，並且具備謀生能力。使社會發展共通的文化價值，能尊重不同，並導引和諧。使國家造就出世界一流的人力，創造國家繁榮。

（二）課程科目

核心基礎科目：英文、數學、科學。
基礎科目：藝術、地理、歷史、現代外語、音樂、體育、科技。

以上十個科目是義務教育中，（所有五至十六歲的兒童都必須學習的內容，現代外語是在十一歲以上才屬必修的）。宗教教育雖未列在十科之內，一般學校也都會教導，但其課程內容是由地方教育當局籌劃，並非屬於國家層次。

（三）課程結構

國定課程的結構包括以下個重要部分：

1. 學習計畫

這是一份非常清楚的說明書，告訴老師到底該教些什麼，所有在此科目中必須教導的內容、技巧、方法、策略等都包含在內。

3. 成就目標

代表學習者在每一個科目中必須獲得的知識、技能與理解，每個目標都有評鑑的工具來記錄學生應該知道什麼，以及能夠做些什麼。

3. 成就水準

每一個成就目標都可以再分為十個水

準，代表由五至十六歲的進步情形。比如level 1相當於一年級可以到達的程度，level 10則相當於中學以上的程度。

4. 集合檔案

每個兒童的成就目標都會記錄下來，當要報告家長時，這些分數都會集合起來如同一份檔案，清楚說明每位兒童的表現。

5. 成就說明

這是針對每一個成就目標及成就水準所做更詳細的說明。

(四) 課程實施

以往由於英國學校制度相當不統一，即使同一地區也常有不同的系統，比如一個十歲的兒童，可能就讀middle school的五年級，也可能是secondary或upper school的三年級，常易混淆。國定課程因此採取關鍵階段的畫定，將五至十六歲的義務教育學習分為四個階段，每一個階段包含二至三年的學習。

表3說明國定課程實施的四個關鍵時期。

表3 國定課程實施的關鍵階段

第一關鍵階段	5歲-7歲
第二關鍵階段	7歲-11歲
第三關鍵階段	11歲-14歲
第四關鍵階段	14歲-16歲

三、國定課程與評量發展

英國的正式國名為「大不列顛與北愛爾蘭聯合王國」(The United Kingdom of Great Britain and Northern Ireland, UK)，全境由靠近歐洲大陸西北部海岸的不列顛群島的大部分島所組成，包括英格蘭(England)、威爾斯(Wales)、蘇格蘭(Scotland)與北愛爾蘭(Northern Ireland)四個地區，隔北海、多佛爾海峽和英吉利海峽同歐洲大陸相望，是一個島國。其各區教育制度差異極大，一般所稱英國教育，係指英格蘭和威爾斯兩區之教育。

英國自1988年教育改革法案通過之後，首次提出「國定課程」(National curriculum)，明定義務教育階段的基本科目、學習方案、成就目標和評鑑設計，其屬於強制性的施行，適用於五-十六歲學生，共計十一年，分為四階段來施行。而其中測驗與評量是國定課程最受重視的部份，每一關鍵階段完成後，需舉行全國性的評量，稱為「中等教育普通證書考試」(General Certificate of Secondary Education, GCSE)。所有國定課程之施行由兩個獨立的行政機關監督分別是英格蘭的學校課程與評量局(The School Curriculum and Assessment Authority, SCAA)及威爾斯的課程與評量局(The Curriculum and Assessment Authority)。

2004年四月由國家評量機構(National Assessment Agency, NAA)全權負責國定課程與GCSE的評量與監督。其主要執掌如下：

1. 確保高品質的國定課程內容。
2. 公開GCSE與相關成就水準的評量具體項目與程序。
3. 採用現代化的評量與嚴謹的審查系統。

- 4.確保評量過程中的公平、公正、公開，並提供一個可供諮詢的場域。
- 5.進行整體評量計畫之評估，並做改進。
- 6.建立教育評量人員的專業知能。

測驗與評量方式是國定課程中最受重視的部份，目前的規定是在每一個關鍵階段完成後，舉行全國性的評量。換言之，全英國七、十一、十四、十六歲的兒童都必須接受測驗。和以往最大的不同是全國性的評量提早進行，以前兒童除接受閱讀測驗外，都是進入高中後才接觸全國性的評量。此種改變的主要動力，來自家長及民意的支持。大多數人都希望儘早得到有關子女學習的回饋（Cutles, 1992）。

全國性評量是一個相當複雜的問題，回顧英國評量發展的歷史，醫學及法律的資格考試始於十九世紀的早期，只有通過入學考試者才有機會進入這些學校就讀。而1944年通過的教育法案雖然在標榜福利國家的政策下，欲使全體國民皆有接受中學教育的機會，不因社經地位的不同而有差別，但法案中認為人的心智能力分為三等，所以應有三種程度不同的學校，依據學生的年齡、能力、性向分類入學。因此，小學畢業後接受十一歲的考試，決定學生進入文法、技術或現代中學。基本上

文法與技術中學，其課程偏重學術性內容且水準甚高而現代中學則側重特定科目的職業教育，因此雙軌精神依然存在。許多學者認為這種十一歲的分流測驗，過早區分兒童為academic與less academic，失之武斷，可能造成難以彌補的錯誤。

依據1988年教育改革法對國定課程之規定，國定課程採取關鍵學習階段的劃定，將五至十六歲的義務教育學習分為四個關鍵階段，每一個階段包含二至四年的學習，如表4所示。為確保每一位學生在義務教育階段的各個學習關鍵期皆能達到成就目標，每一位學生在每一個關鍵學習期結束前必須接受測驗，及接受國定考試（National Tests），經由國家考試來瞭解學生學習成就，以及學校教學表現。

英國義務教育階段共分為四個關鍵學習期，依據英國教育及技能部設計，在第一關鍵期的二年級結束前，必須接受測驗，考試科目為數學與英文；第二、三關鍵期結束前必須再度接受測驗，考試科目是數學、英文與科學；第四關鍵期畢業之前，則可參加屬於資格考試之「中學教育普通證書」（General Certificate of Secondary Education ,GCSE）考試，取得中學學歷文憑，一方面取得繼續升學資

表4 英國義務教育階段關鍵學習之考試類別與科目

關 鍵 期	年 級	年 齡	考 試 類 別	考 試 科 目
第一關鍵期 (key stage1)	一 二	5-6 6-7	SATs (Standard Assessment Tasks)	數學、英文
第二關鍵期 (key stage2)	三 四 五 六	7-8 8-9 9-10 10-11	SATs	數學、英文、科學
第三關鍵期 (key stage3)	七 八 九	11-12 12-13 13-14	SATs	數學、英文、科學
第四關鍵期 (key stage4)	十 十一	14-15 15-16	GCSEs	五至九門GCSE學科

格，另一方面可作為應徵工作之依據。

英國學生成就評量可從法定強制性評量與非法定的選擇評量來區分。法定強制性評量又可區分為教師班級評量與全國性測驗（國家考試）二種。茲就法定強制性評量敘述如下（陳明印，2004）：

（一）教師班級評量

被視為教室裡教和學的一部份，評量部分包括教學方案的所有內容範圍，必須考慮到教室情境中所有學生表現的證據，如經由討論和觀察等評量所獲知；在每一階段的國定課程中，為學生設定每一科的成就標準，教師根據這些標準，透過課堂上對學生的了解，來判斷其學業進步情況；對其進行評鑑，評定每科目的成就，以便提供下一步學習的建議。

（二）全國統一評量

從考試設計來看，英國在義務教育階段的統一考試相當密集，在前三階段以數學、英文、科學為核心考試科目，目的在檢視學生的基本能力，以便作為更進一步學習的基礎；第四階段的考試則在確保義務教育的成效，以及評估學生升學的能力。總言之，國家考試的設計在確保義務教育階段，學生學習成果的品質（National Assessment Agency, 2005）。

肆、OECD/PISA基本能力評量

1999年「國際經濟合作與發展組織」（The Organization for Economic and Cooperation Development，簡稱 OECD）發展跨國的評量學生計畫，此一計畫稱之為「國際性學生評量計畫」（Program for International Student Assessment，簡稱 PISA）。此一計畫推動數年之後，在跨國性的研究架構支持下，在國際已具有相當

的影響力。

一、OECD/PISA內涵

OECD/PISA主要的功能在瞭解個人參與社會活動的能力。評量的內涵包括閱讀能力、數學能力與科學能力三者。

「閱讀能力」在解釋、使用與反應文字的內容，以便達成個人的目標、發展個人的知識與潛能以及社會的參與。閱讀能力的組成與領域面向包括：閱讀不同種類的文字內容，例如連續散文、文告及結構不同文體。表現不同種類的閱讀任務，例如回憶特定的資訊、發展一種解釋或是反應文字的內容。是否能閱讀因不同情境而寫作的文字內容。

「數學能力」在確認、了解與從事數學以及透過數學在世界扮演的角色能作成良好基礎的判斷，並協助個人就目前與未來生活所需，做一個建設性、關懷心與反省性的市民。數學能力的組成與領域面向包括：數學內容--主要數學的「大概念」、模式與問題解決、以及數學在不同情境中的應用。

「科學能力」在結合科學知識與證據本位結論的形成，以及發展相關假設以了解與協助作成自然世界有關的決定，並且透過人類活動來促成改變。科學能力的組成與領域面向包括：科學概念、科學過程技巧以及不同情境的科學應用。整個評量架構的項目如表5（OECD,1999）：

表5 OECD/PISA學力評量的領域

	閱 讀 能 力	數 學 能 力	科 學 能 力
定 義	了解、使用與反應文字的 內容，以便達成個人的目 標、發展個人的知識與潛 能以及社會的參與。	確認、了解與從事數學 以及透過數學扮演的角 色能作成良好基礎的判 斷，並協助個人就目前 與未來生活所需，當一 個建設性、關懷心與反 省性的市民。	結合科學知識與證據本 位結論的形成，以及發 展相關假設以瞭解與協 助作成自然世界有關的 決定。並且透過人類活 動來促成改變。
成 份 及 領 域 的 面 向	<p>閱讀不同種類的文字內容，例如不同形態的連續散文（敘述性）、文告及結構不同文體。</p> <p>表現不同種類的閱讀任務，例如回憶特定的資訊，發展一種解釋或是反應文字的內容或文字的形式。</p> <p>閱讀因不同情境而寫作的文字內容，例如為個人的興趣，或為符合工作的要求。</p>	<p>數學內容包括主要數學的「大概念」，第一測量循環中這些「大概念」以改變與成長以及空間與形狀為主。在未來的測驗循環中，機會、數量的推理、不確定性與依賴關係會應用上去。</p> <p>數學能力即模式、問題解決；這些能力區分為三類：執行程序、連結能力以及數學思考與規約能力。</p> <p>數學在不同情境中的應用，即問題影響個人、社群或全世界。</p>	<p>科學概念包括能源保護、調適、解組分化—從主要的物理、生物、化學等領域選擇、應用這些知識於能源保護、物種保育與物質使用。</p> <p>過程技巧—確認證據、形成、評鑑及溝通結論。這些不依賴先前設定的科學知識，但也不能用在沒有的科學內容。</p> <p>不同情境的科學應用</p>

資料來源：OECD（1999:12）。

二、PISA閱讀能力評量

在OECD/PISA的閱讀能力一詞是一個比較廣義的瞭解。相對的少數年輕人在社會中沒有閱讀技巧，此一計畫所需測量的十五歲學生並不強調他們在閱讀方面技術性的意義，而主要在反應當前的閱讀觀點。此一觀點要學生離開中學時，應該能建構、拓展與反應校內外不同情境所閱讀的連續或不連續的內容。

閱讀與閱讀能力的定義是隨著社會、經濟與文化之改變而不同。學習的概念，尤其是終身學習的特殊性，已經拓展了閱讀能力的知覺與必要性。閱讀能力不再只考慮童年時在學校所獲得的能力，更重要的是把閱讀能力看成是一套知識、技能與策略的進步，個人從生活的不同脈絡與透過同儕的互動中建立這些能力。

認知觀點的閱讀能力強調互動性的閱讀本質與建構性的理解本質（Bruner, 1990; Binkley and Linndkyla, 1997; OECD, 1999）。讀者利用先前的知識和社會與文化共享的文字及情境線索對閱讀的內容產生意義。建構意義時，讀者使用不同的程序、技能與策略去強化、監督與維繫理解。這些過程與策略隨著不同的情境與目的以及不同的閱讀內容而不同。

三、PISA數學能力評量

數學能力（mathematic literacy）這個領域關注的是培養學生去面對未來的挑戰的數學能力。這些能力包括分析、推理及有效的溝通概念，透過形式、公式解決不同領域與情境的數學問題。

在OECD/PISA的數學能力界定中，「能力」（literacy）一詞，在傳統學校的

數學課程中所選擇並強調的數學知識與技能，並不是OECD/PISA所強調的重點。OECD/PISA強調的數學能力重點是數學知識應用在不同的脈絡與不同的方面所需要的反思與洞察能力。「世界」（world）意謂個人所居住的自然、社會與文化的環境。「從事」（engage in）並不是狹隘的侷限於物理或社會行為。「從事」也意味著溝通、採取方式、關聯、評估甚至是欣賞數學。定義中並不限定數學的功能性應用，數學在美學與休閒方面的要素亦包括在此一定義中裡。數學能力定義中的「目前與未來的生活」包括個人生活、職業生活與社會生活，與同儕、親人及社區公民的生活。

四、PISA科學能力評量

在OECD/PISA的科學能力評量方面，以「證明為基礎」的陳述是科學能力之基礎，藉以瞭解測驗觀和理論之合理性。若不能測驗、不是科學理論，充其量只是「偽科學」。

Millar和Oaborne（1998）將現代科學課程視為「閱讀和融合科學和技術資訊之能力」。總結科學能力的不同觀點，科學能力可視為一個階層。Bybee（1997）有四個層級：最低兩個為「名義上科學能力」、「功能能力」、第三為「概念和過程科學能力」、最高則為「多重領域科學能力」。OECD/PISA所提出的科學能力與其第三層級較類似。PISA舉出重要的界定語句，例如：科學能力是使用科學知識界定問題和以「證明為基礎」的結論，瞭解和幫助決定，探究人類活動自然世界和變遷。

OECD/PISA提出三個觀點：科學過程、科學概念、情境。前面兩個觀點是任

務結構和學生表現特徵。第三個觀點將確認評估任務的發展。OECD/PISA架構確認評估的焦點，是基於科學教育的成果。

伍、實施策略分析與比較

本文以英、美與OECD/PISA的評量實施策略為主，比較其間的差異以供參考。

一、美國學生評量的實施策略

NAEP利用由上而下的政策執行方式，進行美國學生基本能力評量的教育政策。展現美國教育領域長期屬於地方分權自治的一個新面貌。美國NCES成立一個獨立專責機構NAGB來專責執行NAEP這一項計畫，並授權此機構來制訂相關教育評量方針與政策。

NAEP有一套標準的評量編制過程，各科目訂有評量設計架構參考依據。

NAEP的評量編製，根據各科評量架構，邀請教育官員、教育專家學者、學校教師、教材專家、與一般公民，決定評量科目的具體目標。NAEP透過這樣協商溝通的管道，審查與修訂各科評量的測驗內容，再將評量的具體目標交由各科的測驗專家進行測驗題庫的建立。最後，各科的評量測驗題庫再一次經過審查、預試、修正、選取樣本預先施測與結果分析，才正式通過決定採用。

評量是整體學生的表現，而非個別學生的成就測驗（assessing groups of students, not individuals）編製評量以專業人員為主導（deciding what is assessed）；

功能擴增下：實施NAEP的州政府與地方教育單位逐漸增加（expanding NAEP's Mission）；各科評量架構的建立

是測驗編製的基礎（beginning with a framework）；

發展與編製評量題目過程嚴謹（developing the assessment questions）；

完成評量小冊子（assembling the assessment booklet）；NAEP的實施方式非全面性評量所有學生，而是選取足夠且具代表性的樣本來參與（selecting the sample）；

鼓勵少數族群學生的參與（accommodating students with disabilities and/or limited-English-proficient students）；

被選取要接受NAEP評量的學生，於接受評量前，其家長有被事先告知的權利（informing parents about NAEP）；

NAEP評量實施的時間約在一月底到三月中，接受NAEP評量應是完全自願的過程（administering and taking the assessment）；

NAEP的評分方式具有信、效度（scoring NAEP）；

評量的報告方式以整體學習成果表現趨勢為目標（reporting student performance）；

評量結果呈現方式多元，可供資料間相互交叉比較（putting results in context）；

評量結果完全公開，有興趣的民眾皆可取得相關資料（presenting results to the public）；建立NAEP專屬網站，提供即時最新消息（NAEP on the Web）；NAEP出版品提供豐富、多元且便利的資訊來源。

二、英國學生評量的實施策略

(一) 教師班級評量

英國的教師班級評量是由老師對兒童各方面表現所進行的持續性評量，逐日或逐週的地記錄下來，成爲一種個人的紀錄卡。老師使用的評鑑依據包括兒童的作業、解決問題的能力、參與團體工作的表現等。依據每一個學生的個別紀錄，老師可以安排適合學生的學習活動，並對兒童所達到的水準予以個別評量。

教師每年需把評量結果匯集成一份學童成就報告，此即爲家長所收到的成績單。教師會根據國定課程的標準，核定學生在學科上所達到的等級。一般而言，七歲的兒童應該達到第二級；十一歲兒童應該達到第四等級；十四歲兒童應該達到第五或第六等級的標準。且此報告需考慮到兒童進步和整個階段的表現，並對照全國等級描述的說明來判斷。

(二) 全國統一評量

由於教師評量會受到教師不同經驗及測驗未經標準化的影響，產生無法比較的歧異，因此尚需透過「標準評量工作」(Standard Assessment Tasks, SATs)，在每個關鍵階段完成後進行全國性標準測驗。

學校將依學生在三個科目中的學習，與國定課程的測驗相比，做出評量。測驗後，學校會向家長提出一份成績單，包含教師評鑑結果、學童在全國測驗中的成績，以及一份與同齡學童成績比較表。十六歲學生所參加任何公定考試(public examination)的成績，亦會登載在成績單上。教師有責任在成績發表後，安排與家長見面，研討學童的學習狀況。

(三) 成績評量的登錄與通知格式

評量的登錄依規定，每年必須送交家

長知悉，稱之爲年度報告(annual report)，此外學校也必須存檔(可供學生轉學之用)，因此檔案的登錄是學校的一項重要工作，由於登錄並無一定格式，所以各校有絕對的自由，但基本上須注意幾項要點：

- 1.紀錄必須包括逐年進行的課程及跨科目活動。
- 2.紀錄必須顯示學校提供廣泛而平衡的課程。
- 3.紀錄必須呈現持續及連續性的說明。
- 4.紀錄必須適合不同的學校轉用。
- 5.紀錄必須顯示整個班級、小組及個人的活動及評量。
- 6.紀錄必須容易處理處理，盡量不增加老師的負擔。
- 7.紀錄必須提供家長、學生及政府有意義的資料。

至於一年一度給家長的年度報告須顯示：

- 1.學校所提供課程的內容介紹。
- 2.各種跨科目活動。
- 3.兒童在各項成就目標或成就水準的表現。
- 4.指出兒童的優缺點。
- 5.指出學習困難的可能原因。
- 6.說明合理的下一步嘗試計畫。
- 7.紀錄個人特別的成就表現。
- 8.肯定兒童的成就及進步。

全國課程學生成就評量通知方式大可分爲二種：一爲，統一通知格式；二爲，非統一通知格式。

統一通知之格式，主要是用在每一關鍵學習階段最後一年，分別依照學生個人、學校和全國三類表現結果，加以統計並通知學生及其家長。每一類的通知書中，必須包含教師評量和全國性評量，每

一學科的整體成就等級。

非統一通知格式，則由地方教育當局或學校自行規範；基本上，格式內容兼具量化記錄和文字描述；量化記錄包括精熟程度、成就等級等，文字描述包括學生學期的學習目標、課程重點、學生表現描述以及優缺點或建議等（陳明印，2004）。

三、OECD/PISA的實施策略

OECD/PISA代表一個新的承諾，靠著OECD國家的政府機構，監控教育系統中學校及學生成就方面的產出。PISA評量的架構及內涵是國際間形成的共識。評量結果每三年出版，預期這些結果將幫助並激發教育改革與學校進步，特別是針對學校或教育系統有相同輸入但結果卻不同的現象，提供檢視的機能。此外OECD/PISA也提供一個國際層級教育系統效能的評量與監控，目前參與的國家已達三十國，是一項頗為值得重視的超大型研究計畫，尤其是許多寶貴資訊的正確解讀，隱含各種時、空的意義。

OECD/PISA包含多元評量的範疇，三年一循環補充最新資料，以達到定期的監督學生進步，進而符合關鍵學習的目標。評量對象的年齡群為十五歲，於義務教育接近結束時進行評估，提供教育系統有用的指標。大部分OECD國家的教育超過十五歲，在這個年齡前都有一個較共同的課程，就此基礎進行的相關知識與技能。知識與技能評量的定義並不完全以國定課程為主。

PISA主要考量未來的生活。學校所學的技能與生活所需的技能無法明確區分，但是瞭解造成這種區分的存在是非常重要的。OECD/PISA不排除學校基礎課程的知

識與瞭解，但是評量想瞭解的是更廣的概念與技能的獲得，及這些知識的應用。

陸、結論與建議

本研究探討的主題包括：1. 各主要國家學生基本能力評量的內涵與優缺點；2. 各主要國家推動學生基本能力評量的策略；3. 這些學生基本能力評量的呈現方式有何借鏡之處。我國已進行國中階段的「基本學力測驗」，未來如果要全面推動「學生基本能力評量」，可能的配套措施為何？本研究的主要結論與建議如下：

一、結論

（一）美國學生基本能力評量之優缺點

上述NAEP內涵之探討進一步呈現美國學生基本能力評量的優缺點及特色。美國的學生基本能力評量具有下列優點：

1. 獨立行政組織執行政策；
2. 標準參照測驗；
3. 由上而下強制性評量；
4. 教育經費補助的依據；
5. 各科評量時間分配適宜，不讓學校過分負擔。

美國學生基本能力評量的缺點包含：

1. 無國定課程標準可遵循；
2. 地方教師反彈；
3. 評量題目設計的缺失。

（二）美國的發展經驗可供參考之處

1. 長期性的執行與規劃；
2. 閱讀與數學為重點評量科目，國際的比較性評量為主導；
3. 建立學生能力評量的資料題庫；
4. 評量題型的架構使長期教育發展趨勢的評量結果可以互相比較；

5. 專責單位的設立使學生基本能力的掌握持續不間斷；
6. 重點科目加強評量的密集性；
7. 學生評量結果的呈現方式有所區分，主要評量（Main NAEP）呈現學生成就的結果，長期發展趨勢評量（Long-term Trend NAEP）提供全國性資料而非學生的個別成就；
8. 學生基本能力評量的各科目架構，每十年重新檢視更新；
9. 學生樣本採取抽樣方式，其中包含少數族群的學生代表；
10. 積極參與掌控學生基本能力的學區可獲較高教育經費的支持。

（三）英國學生基本能力評量之優缺點

英國發展的評量系統具有下列的優點：

1. 賦予人民應有的權利

國定課程為所有的兒童，不論是出生、文化、種族、性別、階級、能力及障礙者，界定各種完善的學習領域，促進教育機會均等，並賦予每個人民應有的學習權利，期以發展自我，進而自我超越。

2. 建立一個全國性的標準

國定課程對學生、家長、教師、政府和社會大眾，清楚設定出學生和成就目標的期望。而且對所有學生，在學科中的應有表現，制訂出全國性的標準。這些標準可作為制訂、改進、測量、達成這些成就的指標，甚至作為監測的評核，以比較個人、團體和學校間表現的參考依據，以了解學生在學習過程中的程度與差異性。

3. 促進連續性和一致性

國定課程引導一個全國義務教育一致性架構的產生。該架構促進課程的連續性，且以充分彈性以確保學生學習的進

步；它也會強化學生在各教育階段和各學校的轉換，並提供個終身學習的基礎。

4. 促進公共大眾的理解

國定課程一致性的課程標準會增進社會大眾對學校教學成效，特別是對義務教育成果的理解。尤其也可藉此提供非教育界人士以及教育專業團體間，包括學生、教師、雇主間，發展出教育議題討論的共同討論基礎。

英國在評量實施過程所面臨的問題或缺點方面包括：

1. 教師不願配合的困擾

教師除擔任日常教學外，須依據國定課程要求，必須隨時掌握學生在每一科目的每項成就目標及水準的表現，並加以錄。此外還必須配合國定課程的評量，進行標準化測驗，更使老師疲於奔命。加上測驗結果常與老師評量相反，無法幫助老師作出有意義的決定。

2. 信度、效度的問題

- （1）試題方面：當使用相同的標準或量尺，則無法掌握標準以外的品質；
- （2）學生方面：學生或學校間因其主客觀條件不一，使用同一標準加以比較，甚至將其結果公布，實在有失公平；
- （3）教師方面：無法保證教師評量完全客觀。

3. 評量領導教學的問題

教師專注於學生是否能在考試中得到高分，則會影響其教學方式與內容；學習若只為了應付考試，則知識常被切割為學習零碎的知識；全國性評量如果成為主流，則教師平日的形成性評量功效盡失。

（四）英國的發展經驗可供參考之處

英國國定課程評量是希望改變以往偏

頗的觀念，藉由教師持續紀錄兒童發展及標準化的國家測驗，提供兒童表現的可靠資訊。此種評量策略結合形成性、診斷性、綜合性與評鑑性的功能，具以下特點：

1. 確認學生已經達到的目標，幫助他們計畫下一個學習目標，也稱為「形成性」的評量；
2. 顯示兒童學習的缺點及困難，幫助其有效克服，也稱為「診斷性」的評量；
3. 給與老師、家長、兒童一個全面性的綜合報告，在固定時間報導學生到達的程度，也稱為「綜合性」的評量；
4. 提供家長、老師、教育當局有關學校表現績效，用以比較學校之間的辦學差異，也稱為「評鑑性」的評量。

（五）OECD 學生基本能力評量之優缺點

OECD/PISA 評量的優點如下：

1. 強調教育系統的評量，可以做為教育系統運作成效的診斷；
2. 評量內涵以生活化能力、終生學習技能的養成為主；
3. 評量具有國際比較的功能，可讓參與國家互相觀摩比較；
4. 評量提供時間序列的比較功能，從2000年至2006年，其後也已規劃後續的執行計畫；
5. 評量結構嚴謹，評量結果的資訊以專書形式出版，相當透明化；
6. 評量結果資訊的呈現除了顯示平均成績之外，也呈現學生表現的離散程度，具有多元呈現的功能。
7. 評量的結果對各國的教育政策有相當的影響力。

OECD/PISA 評量的問題或缺點包含：

1. 跨國使用的試題，有翻譯上及文化差異上的困擾；
2. 非全面性的評量，閱讀能力、數學能力與科學能力有個別的評量重點，並每隔三年一次，週期有點鬆散；
3. 以抽樣方式進行且侷限義務教育終了階段的評量較難達成及時的診斷功能。

（六）OECD/PISA的發展經驗可供參考之處

1. OECD/PISA是跨國合作、比較的典範，所強調的是當前最先進的教育理念，具相當參考價值。我國於2004年加入此一評量計畫，OECD/PISA 評量的運作模式、國際的研討互動，對國內提升學生素質的相關措施有積極性意義；
2. OECD/PISA採取各主要國家分工合作發展的模式，參與OECD/PISA 可以與國際接軌，除了有助於國內學者的研究視野之外，也提升台灣的國際能見度。

二、建議

本研究以美國NAPE、英國國家課程評量及OECD/PISA為對象，分別探討其學生評量內涵、評量策略、及可供參考之優劣點。綜合分析結果，本研究對於未來國內發展全面性基本學力評量有如下建議：

（一）評量機構方面

1. 慎選發展評量的機構，做長期性、系統性的研究發展；
2. 採取以校際或是機構分工的設計模式，結合學者專家的專長，可以快速達到分工整合的功效；

（二）評量內涵方面

1. 審慎參考國際的評量架構與評量內涵，做為發展國內的評量系統之參考；

2. 打破以教科書本為範疇的評量型態，以學習技能及生活技能為評量學生表現的重點；
3. 建立長期性評量題庫的建置、發展單位。

(三) 評量實施方面

1. 初期先以抽樣的方式進行試驗，以建立基本的評量內涵；
2. 最終以建立全國性的學生評量資料庫為目標。

(四) 評量結果的呈現

1. 可以顯現全國性學生表現的資訊；
2. 可以呈現區域比較及校際比較的資訊；
3. 評量結果的呈現具有時間序列的意義。

過去台灣師大發展的「基本學力測驗中心」，以國中基本學力測驗為主，目的在做為學生升學之選擇與分發的依據，與國際間發展的基本能力評量工具不同。近年來，國立教育研究院籌備處亦著手進行國內「基本學力評量」的研發工作，惟目前只是研究階段，嘗試建立基本的測量題庫，距離全面推動仍有一段時間。就現階段的發展來看，重新思考國內全面進行基本學力評量的發展策略有其必要。

為今之計，建議教育部先成立「學生基本能力評量任務小組」，配合教育部的政策，就此項任務推動的時程，重新擬定發展計畫，並參考前述建議規劃行動策略。

參考文獻

- 陳明印 (2004)。英國國定課程學生成就評量制度。研習資訊，21 (6)，90-108。
- Binkley, M., & Linndkyla, P. (1997). Teaching reading in the United States and Finland. In M. Binkley, K. Rust and T. Williams (Eds.), *Reading literacy in an international perspective*. Washington, D.C: U.S.Department of Education
- Bruner, J. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.
- Bybee, R. W. (1997). Towards an understanding of scientific literacy. In W. Grabe and C. Bolte (Eds.), *Scientific literacy: An international symposium*. IPN, Kiel.
- Cutler, B. (1992). How is your child doing at school? *British Journal of Curriculum & Assessment*, 3 (1), 21-23.
- Department for Education and Employment (1996). *Results of the 1996 National Curriculum Assessment of 14 year olds in England*. London: The Author.
- National Assessment Agency (2005). *National curriculum tests*. Retrieved October 15,2005, from <http://www.naa.org.uk/tests/index.aspx>
- Millar, R., & Osborne, J. (1998). *Beyond 2000: Science education for the future*. London: King's College London School of Education.
- National Assessment Agency (2005a). *Key stage 1 (ages 5-7)*. Retrieved October 15,2005, from http://www.qca.org.uk/12818_13017.html
- National Assessment Agency (2005b). *Key stage 2 (ages 7-11)*. Retrieved October 15,2005, from http://www.qca.org.uk/12818_13016.html
- National Assessment Agency (2005c). *Key stage 3 (ages 11-14)*. Retrieved October 15,2005, from http://www.qca.org.uk/12818_13015.html
- National Assessment Agency (2005d). *At a glance*. Retrieved October 15,2005,

- from http://www.qca.org.uk/12818_13011.html
- National Center for Education Statistics, National Assessment of Educational Progress (NAEP)(2005a). *2005 mathematics assessment: Comparison of the 2005 and 1990-2003 mathematics frameworks*. Retrieved October 5, 2005, from <http://nces.ed.gov/nationsreportcard/mathematics/frameworkcomparison.asp>
- National Center for Education Statistics, National Assessment of Educational Progress (NAEP)(2005b). *2005 mathematics assessment: Distribution of mathematics questions 2005*. Retrieved October 5, 2005, from <http://nces.ed.gov/nationsreportcard/mathematics/distributequest2005.asp>
- National Center for Education Statistics, National Assessment of Educational Progress (NAEP)(2005c). *1996 and 2000 science assessments*. Retrieved October 5, 2005, from <http://nces.ed.gov/nationsreportcard/science/timedevoted.asp>
- OECD (1999). *Measuring student knowledge and skills: A new framework for assessment*. Paris: Author.
- The Nation's Report Card (2005a). *Overview of the Nation's Report Card*. Retrieved October 5, 2005, from <http://nces.ed.gov/nationsreportcard/about/>
- The Nation's Report Card (2005b). *What are the differences between long-term trend NAEP and main NAEP*. Retrieved October 5, 2005, from http://nces.ed.gov/nationsreportcard/about/ltr_main_diff.asp
- The Nation's Report Card (2005c). *Schedule for the State and National Assessment of Educational Progress (NAEP) from 2005-2017*. Retrieved October 5, 2005, from <http://nces.ed.gov/nationsreportcard/about/assessmentsched.asp>
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) (2005). *2005 reading assessment: What does the NAEP reading assessment measure?* Retrieved October 5, 2005, from <http://nces.ed.gov/nationsreportcard/reading/whatmeasure.asp>
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) (2005). *2005 reading assessment: Distribution of reading question 2005*. Retrieved October 5, 2005, from <http://nces.ed.gov/nationsreportcard/reading/distributequest.asp>
- U.S. Department of Education(2005). *U.S. department of education principal office functional statements*. Retrieved October 5, 2005, from http://www.ed.gov/about/offices/list/om/fs_po/osods/os_orgchart.html

註釋

本研究改寫自國立教育資料館九十四年度專案研究報告—「基本能力測量跨國發展經驗之比較研究」，特向國立教育資料館致謝

