

台灣學生學習成就評量資料庫建置計畫之探究

陳清溪/國立教育研究院籌備處主任秘書

林欣怡/國立教育研究院籌備處約聘幹事

壹、前言

行政院教育改革總諮議報告書建議：針對學生的基本學力進行研究，編製標準化評量工具，進行學生基本學力鑑定，以便及早提供補救教學，帶好每一位學生（行政院教育改革審議委員會，1996）。國民中小學九年一貫課程實施後，部分民意代表、學者或教師認為學生能力降低，尤其是建構式數學，有些人認為學生數學能力提升，有些人認為降低，造成正反雙方支持者的爭辯，就是缺乏具客觀、科學的證據。因此，研發標準化成就測驗工具，針對台灣中小學生學習表現情形，進行全國性的檢測工作，以建立長期性、整體性的學生學習成就資料庫，以供課程與教學決策之參考，是刻不容緩的工作。

事實上，我國中央或地方教育當局及學者專家，都曾經針對某個年級的某些科目，進行學生學習成就的評量工作，但卻缺乏長期、跨科目、跨年級的整合型長期追蹤評量，因此，學生的學習成就資料不盡完整，自然也就無法客觀的比較學生在學習方面的長期變遷趨勢。自實施九年一貫課程後，台灣學生的學習成就到底是提升或下降，一直受到爭議，似乎沒有研究團隊可以提出數據來解答這個疑惑，關鍵就在於我們欠缺一個

研究團隊，致力於建置全國性的學生學習成就資料庫，以便縱貫性比較學生學習成就。

教育基本法（1999）第九條明定中央政府之教育權限，該條第六項為：「教育統計、評鑑與政策研究」。因此，透過評量來瞭解學生在各學習領域（科目）的表現，以評鑑學生的學習優劣，乃中央政府之教育權限之一。九十三年四月二十日教育部國民中小學九年一貫課程推動工作小組第四十三次會議決議：「有關學生學習成就調查是本部研訂課程與教學政策之重要參據，有必要進行常態性之資料建立。教育部乃於九十三年五月十九日以台國字第 九三 六六七三三號函，請國立教育研究院籌備處（以下簡稱本處）針對國民中小學學生學習成就建立常態性之資料庫，俾為本部研訂課程與教學政策之重要參據。本處旋即規劃台灣學生學習成就評量資料庫（Database Of Taiwan Student Achievement, 簡稱DOTSA）之建置計畫，業經教育部於九十三年十月十一日以台國字第 0 九三 0 一二七四九六號函備查。

貳、教育指標與調查方法

在探討國內外大型學習成就評量資料庫計畫前，先探討「教育指標」的意義，接著

簡述縱貫性與橫切面調查方法之利與弊。

(一) 教育指標與教育指標系統 (Education indicator vs. Educational Indicator System)

談到教育，總避不開所謂的「能力指標」、「學力指標」、「教育指標」……等等，這些詞彙中都包括了「指標」(Indicator)的概念，究竟在教育領域中「指標」的意義為何？其操作型定義(operational definition)又該如何詮釋？在進一步探討全國性或國際性的學生學習成就評量計畫前，宜先釐清「指標」的意義為何？依據 Shavelson, McDonnell, and Rand (1991) 的闡述，指標是指一個別的(individual)或組合(composite)的統計量(statistics)，亦是教育的基本架構(construct)之一，且得以在政策脈絡(policy context)下發揮效用。吳清山與林天佑(1999)指出教育指標的意義其實可由二方面來詮釋：(一)教育指標是評估教育體制及政策運作的標竿，(二)教育指標是導引日後教育體制運作及教育政策考量的重要依據。

因此，單一指標是很難在一複雜脈絡中，提供充足有用的資訊。以學校本位管理(school-based management)來舉例，定義學校本位管理似乎不可能單靠一個指標，因為學校本位管理其實包含了教師增能賦權(teacher empowerment)、責任績效、家長的參與程度等核心指標。另外，倘若要定義學校的社經地位，若僅以單一指標，如符合社會補助低收入戶在一學校中的人數比例，

似乎不夠周全，但如果將學生家長的年收入、職業、最高學歷等因素同時併入考量，反而較能確實反應學校社經地位所代表的概念。因此，教育指標系統(educational indicator system)因應而生，可針對某一環境情況，彙集各單一指標成為指標系統，而在這一指標系統裏，各個單一指標代表不同層面的內涵，以均等或非均等的比重(loading)組成一指標系統，而提供更正確可靠的資訊。

在1983年4月，美國國家教育卓越促進委員會(National Commission on Excellence in Education)提出「危機中的國家」(A Nation at Risk)報告書，這份報告書客觀不偏頗的陳述美國當時教育的現況，說明當時美國教育下學生學習成就大幅下降且落後其他工業發展中國家七倍以上，因而轉向致力於學生基本學科教育。「危機中的國家」報告書歷歷指出學生在學習成就上危機的指標，這些意味著教育危機的指標簡述如下(National Commission of Excellence in Education, 1983)：

1. 有二千三百萬的美國成人仍是文盲(functionally illiterate)。
2. 有13%的美國17歲學生仍是文盲，在少數民族(minority)的青少年文盲比例更高達40%。
3. 美國17歲學生在科學成就測驗方面的成績與1969, 1973, 及 1977年的國家科學評量相較之下，呈現持續走下坡的趨勢。

4. 許多美國17歲學生缺乏社會所期望的高階學術能力 (higher order intellectual skills)；其中約有40%缺乏從書面內容推論的能力；其中只有五分之一有能力寫具有說服性的文章；而只有三分之一能解只需幾步驟的數學問題。
5. 美國高中學生在大部分的標準化測驗的平均成績較26年前當Sputnik實施時低落。
6. 超過一半的資優學生的能力與他們在學校的學習成就不相符合。
7. 美國大學入學考試 (College Board's Scholastic Aptitude Tests, SAT) 成績自1963到1980 持續下滑；平均字彙成績下滑了50分，而數學平均成績也下滑了將近40分。
8. 大學委員會的成就測驗同樣發現到幾年來物理和英文成績也成穩定下降。
9. 在美國大學入學考試 (SAT)有優異成績表現 (高於總分650分或以上)的學生不管在數量上和比例上也呈現大幅度的下降。
10. 在1975 與 1980年之間，公立四年制的大專院校裡的數學補救教學課程增加了百分之72，這些補救教學課程佔了大專院校裡數學課程的四分之一。
11. 大學畢業的學生平均學術成績也較之前低落。
12. 商業界及軍方的領導者抱怨，他們花費數百萬美元在花費龐大的補救教育

及職業訓練上，僅為了閱讀、寫作、拼音、及計算等基本能力上。倘若沒有這些補救工作，這些員工或軍人根本無法開始或勝任現今社會所需具備的能力與技術。

這些震驚美國社會的指標歷歷在目，決策者、專家學者、家長憂心忡忡，莘莘學子到底學了沒有，又學到了什麼？國家所培育的人力資源素質，竟然無法配合工商社會前進的腳步。反觀台灣的教育似乎也同樣需要這樣的一份報告書，客觀定位台灣的教育政策方針與反映到的學生學習狀況，特別是當下教育改革促成教育制度的放鬆，九年一貫課程的實施，一綱多本等等，我們實在應該全面性、整體性，來了解我們的學生到底知道些什麼。

國內現行九年一貫課程係以能力指標的意涵來詮釋學生的學習表現，教科書亦根據能力指標來編寫教科書，造成各家出版社對能力指標解讀不同，形成教科書一綱多本，各版本內容不一的現象，因此如何界定能力指標的意涵，依據能力指標來編製成就測驗工具，以評量學生的學習表現，是必須克服的問題。

(二) 縱貫性 vs. 橫切面調查方法 (Longitudinal vs. Cross-sectional survey approach)

任何一個研究都因依其研究目的和研究對象的本質，來決定研究的設計方式。研究的設計依資料蒐集的次數或時間長度主要分為二類 (王文科、王智弘, 2004)：縱貫式調

查方法 (longitudinal survey)與橫斷式調查方法 (cross-section survey)。橫斷式調查係指對母群體 (target population)所選取的樣本僅進行調查一次，可基於研究目的了解各變項之間的因果關係或互為因果關係，及各組間的差異。一般來說，此類調查適用於較大母群體，所以大多採用分層隨機抽樣或簡單隨機抽樣。但由於橫斷式調查只對樣本 (sample)調查一次，所以較無法探視研究對象的縱向變化差異，此為橫斷式調查之限制。

縱貫式調查係指對一母群體所選取的樣本進行多次調查測量，時間之間隔，視研究內容及目的，可長可短，縱貫式調查可延續好幾年。若依其研究對象，縱貫式調查可歸納為三種：(1) 趨勢研究 (trend study)；(2) 小組研究 (panel study)；(3) 同群期研究 (cohort study)。

趨勢研究主要是針對一母群體進行一段時間多測點調查測量，每一測點為一般母群體 (general population) 中所抽測的不同樣本，也就是說，在不同時間點選取相同母群體裏的不同樣本。

小組研究主要是針對一母群體中的相同樣本進行一段時間多測點調查測量，此研究調查方法因只對一相同樣本蒐集資料，而避免了不同樣本所可能伴隨的外在變項 (extraneous variables)，也能清楚測量時間因素下的變異量 (net difference)。但小組研究調查的主要限制在於樣本耗損 (sample attrition)，因為樣本人數可能研究問題之爭

議性或間隔時間太長而流失，所以在小組研究調查的施測間隔不宜太長。

同群期研究在於針對特定母群體中的不同樣本進行長時間追蹤研究。所謂的特定母群體 (specific population) 不同於一般母群體 (general population)，舉例來說，某研究者致力於青少年批判性思考能力的培養訓練，在研究調查設計上，可在特定研究期間內的一時間點對都會區國中一年級學生的母群體抽樣；三年後，再對都會區高中一年級學生的母群體抽選樣本，也就是說，不同於上述兩種研究都只針對有一個一般母群體進行抽樣及資料蒐集，同群期研究事實上牽涉到至少二個母群體 (例如，特定母群體)，此種方法彌補了小組研究中樣本耗損 (sample attrition) 的限制。

各種調查方法，各有其利弊。研究者在規劃任何一項研究調查時，須審慎考量調查設計的方法是否能回應研究目的、回答研究問題，研究結果亦能概括母群體。調查的樣本數量及特性是否能概括 (generalize) 代表母群體 (target population)，往往牽動研究的外在效度 (external validity)。

參、國內目前建置教育資料庫的經驗

國內正建置中的教育資料庫主要有：(一)「台灣教育長期追蹤資料庫 (TEPS)」，主要在蒐集全國中等學校學生在身心發展、學習活動、學習成就表現及其影響因素的多

面向資料(張荳雲,2003);(二)「台灣高等教育資料庫之建置及相關議題之探討」,在建立有關全國高等教育校院特質與狀況、課程、經費、校園環境設備、校園文化、各科系在校學生素質、學習行為與表現、各學門畢業生、以及教職員素質等資料的全國性資料庫(彭森明,2003)。

肆、他山之石可以攻錯

建置一完整且客觀的全國性學生學習成就資料庫,是刻不容緩的工作,以縱向瞭解學生的學習表現,特別是在九年一貫課程實施之後,作為日後決策之參據。隨著國際互動日益密切,與其他國家(如美國、日本、英國)的接軌比較更應列入短程目標。事實上,美國對於評量該國學生學習成就及國際性的比較接軌一直不遺餘力,且頗有成就,值得我國借鏡效法,所以本文將簡略介紹四個國際性的有關學生成就評量的整合型研究計畫。

一、國際數理趨勢研究(The Trends in International Mathematics and Science Study, TIMSS)

就以美國的國際數理趨勢研究(TIMSS)來說,此趨勢研究為因應美國教育學界對於建立美國學生在數學、科學方面學習成就定位及希望與其它國家學生在這二個領域比較的強烈要求。TIMSS在數理方面最嚴謹且容

易理解的評量研究。就目前為止,TIMSS分別在1995年、1999年、2003年進行評量。美國各州可視其意願及需要,自由加入此比較研究行列。美國各州對於此研究計畫之參與並非強迫性,事實上,從過去的10年裏,美國各州已逐步意識到學生在學習數理方面國際定位的重要性,因而趨之若鶩的參與此研究計畫(National Center for Education Statistics,2002)。

二、國家教育進展評量(National Assessment of Educational Progress, NAEP)

國家教育進展評量(NAEP)由美國教育部國家教育統計中心(National Center for Education Statistics,NCES)統籌規劃,由下設的教育統計委員會(Commissioner of Education Statistics)負責依法執行,另外,由美國國會任命但不隸屬於教育部的國家評量委員會(The National Assessment Governing Board, NAGB)則負責制定相關政策,研發評量大綱(framework),施測規準(test specifications)等。簡言之,國家評量委員會負責研發規劃,而教育統計委員會則負責實施執行此長期追蹤研究計畫。

國家教育進展評量(NAEP)例行每四年施測一次(2004年、2008年、2012年),針對四年級(九歲)、八年級(十三歲)的閱讀及數學領域作長期追蹤研究(trend study)。該目的在提供有關美國青少年在基本成就學習上縱向

變化的量化依據。此外，除了四年一輪的追蹤研究，國家教育進展評量更以一案多軌的方式同時規劃全國性(national)及各州(state)評量，以反應回饋教學及教育體制之現況。全國性及州內的評量，以兩年為一評量輪次，共評量閱讀、數學、科學等三個學科領域，並預期於第四評量輪次，也就是2011年時替換科學領域，加入寫作能力領域，此外，非主要基本能力學科如公民學、藝術、美國歷史、世界歷史、地理、經濟及外國語言等，則擇年實施(National Center for Education Statistics,2002)。

三、國際學生評量(Program for International Student Assessment, PISA)

國際學生評量(PISA)為一跨國際性研究專案(multi-national project)，由世界經濟合作與發展組織(Organization for Economic Cooperation and Development, OECD)主導執行。以美國為例，教育部國家教育統計中心提供經費，負責美國國內的研究調查，而資料收集和分析則委外給WESTAT公司負責。PISA為近來教育指標系統(Educational indicator system)的代表，其目的主要在(一)評量學生能否將其所學的課外知識與技能應用至日常生活的準備程度；(二)反應現今教育體制及社會對於15歲青少年的影響；(三)象徵國際互動進而提供資料給世界各國的決策立法者與研究學者。

國際學生評量(PISA)自2000起每三年評量一次，針對15歲學生的閱讀(reading literacy)、數學(mathematics literacy)和科學(science literacy)三個主要學科領域，以三年為一評量輪次，PISA雖然於特定年同時評量閱讀、數學、科學，但每一學科領域以不同比重分配。2000年主要著重在閱讀，約佔評量時間的三分之二，而數學與科學則合佔三分之一。同理，PISA 2003年則著重在數學領域，包含問題解決能力(problem-solving skills)，2006年則著重在科學領域，包含資訊及溝通技能的運用，並將嘗試運用電腦科技輔助評量(computer-assisted assessment)。簡略來說，PISA評量結果是以0-1000分為給分範圍，平均分數為500分，標準差為100分，大部分的分數落在200至800分之間。又每一學科領域分為五個達成層次(proficiency levels)，受測學生需答對一層次中的大多數題目，才視為精通熟練此一能力。

PISA之所以選擇15歲青少年學生為特定單一研究對象，其背後考量有二。PISA研究團隊認為15年歲的學生正值切入成人生活的轉換點，所具備的不單單限於閱讀能力，了解特定數學公式，或科學概念，更要能應用所學的知識技能，解決日常生活中所遇到的各種情況。再者，針對長期追蹤15歲學生，可使參與研究的國家，比較其莘莘學子學習成果，反映在社會及教育體制的影響，特別是學生在完成義務教育之後，是否已做好面對成人生活的準備。另外，對於參與此研究

計畫的國家中的立法決策者，不單單想要獲知同時各國學生學習成果的差異，他們更有興趣的是針對某些特定族群的學習成就，例如父母的教育程度(Parent's education)，父母的社經地位(socio-economic status)，國籍(national origin)，及性別等等。例如，在2000年所有參與研究計畫的國家中包括美國，15歲女生在閱讀領域方面的表現遠勝過15歲男生，而在數學及科學領域方面，男女生的表現則為相當(National Center for Education Statistics, 2002)。

四、美國教育與其他八國的比較指標 (Comparative Indicators of Education in the United States and Other G8 Countries)

G8 (Group of Eight) 計畫實為一大型整合資料庫，運用整合國家既存的相關研究，其主要資料來源有三 (National Center of Educational Statistics, 2002)：(一)為「全國教育系統指標」(Indicators of National Education System, INES)，(二)為「教育成就評量國際協會」(International Association for the Evaluation of Achievement, IEA)，包括了1999年的國際數理趨勢研究(TIMSS, The Trends in International Mathematics and Science Study)，(三)為「國際學生評量計畫」(2000, 2003)(Program of International Student Assessment, PISA)。

以G8計畫來論，世界上的前八大工業化國家—加拿大、法國、德國、義大利、日本、蘇聯、英國、及美國均加入此比較研究行列。美國的G8計畫之所以選擇其他七大國家，主要是因為這些國家與美國在經濟上的發展程度非常類似，而且這些國家亦是美國經濟發展上的主要競爭對手。事實上，利用相似經濟條件的國家來比較其學生在學習方面的定位，較能確保在此研究比較上基準點(baseline)的合理性。另外，為確保跨國性比較的整合性(Integration)及可實施性(Accessibility)，G8計畫要求參與此研究計畫的國家需依據美國在1997年研擬的「國際標準教育分類」(International Standard Classification of Education, ISCED)定義，來各自建置國家的基本教育資料。如此進行國際性的比較研究，的確需要考量到各國在教育體制上的不一致，若無考慮教育體制的差異而逕行比較，必將嚴重影響到研究結果的信效度。

G8計畫列出共31個指標，而這31個指標林林總總代表了美國及其他G8國家在教育體制上的顯著差異面，大致上，G8計畫報告區分這31個重要指標為以下五大類指標群：(1)教育人口統計與社會經濟背景，(2)學前教育(Preprimary education)與初等教育(Primary education)，(3)中等教育(Secondary education)，(4)高等教育(Higher education)，(5)教育界與勞工界(Education and the Labor force) (Sherman, Honegger, & McGiven,

2003)。第一類的教育人口統計與社會經濟背景指標群可定義為國家全體人民的教育成就，其中涵蓋了個人一生中的教育註冊(Enrollment in education)指標，與初等、中等及高等教育中的公私立學校學生註冊的分布情況(Distribution of enrollment)。而第二、三、四類指標群，統計各教育階段的學生註冊率，再者調查學生-老師比例(pupil-teacher ratios)、老師教學策略及學生對數學、科學等學科的學習反應與態度，並比較學生學習成果(student outcome)，其中包含了國際評量成績和體制成效(system outcome)如畢業率。此外，第二、三、四類指標群另外涵括各國財政上投資在教育方面的比較。最後，第五類指標群意欲了解教育成效(education attainment)與勞工界投入參與及所收入(earning)之相互關係(National Center of Educational Statistics, 2002)。由上述可知，G8 計畫所涵蓋的四個教育程度由學前教育(Preprimary education)，初等教育(Primary education)，中等教育(Secondary education)一直延展到高等教育(Higher education)。此研究統整其他相關研究，以其具前瞻性、審慎度，確實值得國內參考借鏡(National Center for Education Statistics, 2002)。

伍、簡介「台灣學生學習成就評量資料庫(DOTSA)」研究計畫

本處在接到教育部的公文指示後，立即著手規劃建置台灣學生學習成就評量資料庫事宜，邀請相關單位及人員，召開學生學習成就評量資料庫建置諮詢會議，歷經三次學者專家團體諮詢會議，一次國中基本學力測驗推動工作委員會諮詢會議，及彭森明教授、張苙雲教授個別訪談會議，彙集多次學者專家諮詢意見，並參考國內外有關學生學習成就評量資料庫建置計畫，擬定「台灣學生學習成就評量資料庫」研究計畫乙份函報教育部備查。

國內學者專家或各縣市政府教育局，都曾經針對某些年級的某些科目，進行學生學習表現的評量工作，但大都缺乏標準化的測驗工具，也缺乏長期性、系統性、客觀的量化資料，無法對學生的學習成就作整體的分析比較，不利於課程與教學政策的訂定，因此，國立教育研究院籌備處乃結合國內學者專家、各縣市政府教育局的力量，研擬計畫開始建置學生學習成就評量資料庫。

一、本資料庫的目的

(一) 建立國民中小學及高中職學生學習成就長期資料庫，以追蹤、分析學生在學習上變遷之趨勢，進而檢視目前國家教育體制與政策實施之成效。

(二) 瞭解國內學校教育之現況，作為教育部研訂課程與教學政策，縣市政府教育局及學校推動補救教學之重要參據。

(三) 分析學生在學習成就上之表現差異，藉此評估學生未來在學術方面能力之發

展與社會期許。

(四) 提供完整、標準化的量化資料，以供國內外相關研究人員深入探討學生學習成就方面的相關政策議題。

(五) 建立本國學生學習成就評量資料庫，同時考慮與國際接軌，利於加入國際比較行列，藉以瞭解台灣教育之獨特面與優缺點。

二、本資料庫之內容

(一) 研究設計

本資料庫是採整體規劃，分年建置的方式逐漸充實各學科評量資料。本資料庫之研究樣本以趨勢研究(trend study)為主，對國小、國中、高中職等三個教育階段進行施測。然而為更進一步瞭解其環境脈絡(context)，藉以提昇施測結果分析之全面性(completeness)，取小部份具代表性的樣本進行追蹤小組研究(panel study)。資料庫內容主要包含國語、英語、數學、自然科學、社會等五個學科領域，並視政策需要，再增加其他學科領域，以分年建置方式逐步完成。測驗內涵包括認知、技能、情意三大層面，初步規劃以認知層面為主，技能及情意部分則採小規模方式研發。

資料庫內容主要在收集學生學習成就表現方面的資料，對於影響學生學習成就之相關因素，亦將同時進行資料的蒐集。而影響學生學習成就之相關因素，大致可歸納為五大類：(1)學生個人基本資料，包括性別、居住地、在校學期成績、喜歡科目、交友情

形、課餘的學習情形(如補習、上才藝班)；(2)學生家庭資料，包括家庭結構、家庭經濟，父母教育程度、婚姻狀況、社經地位、父母對子女的教養態度、教育期望及家庭讀書、生活環境；(3)班級結構，包括人數、編班情形、班級氣氛；(4)教師的教學方式與策略、教學態度、與學生相處情形；及(5)學校規模等。這些相關因素可供日後分析資料時，能更深一層探討學生學習成長之變遷因素，作為補救教學及改善學習環境之參考。

(二) 抽樣設計

本研究將採層級隨機抽樣(multi-staged, stratified random sampling)，因層級隨機抽樣可以確保所施測之樣本具有全國代表性。為確保本資料庫所抽取之樣本具有全國代表性，且因台灣各縣市人口多寡各異，為充分顯現教改後台灣學生學習成就實際情形，故採二階段隨機抽樣設計。第一階段採分層叢集隨機抽樣，根據縣市、人口密度、及班級數等三個變項進行分層。第二階段則根據所抽取到之樣本學校對每一層以學生個人為單位，進行簡單隨機抽樣。本研究以全國國小四年級、六年級、國中二年級及高中職二年級學生為施測年級。為進行長期性追蹤研究，九十五年受測之國小六年級學生中的10%，將列為長期追蹤對象。每兩年追蹤一次，直到高中職學校畢業為止。

(三) 研究工具

本研究將參考國民中小學九年一貫課程綱要、高中職課程綱要內涵、各版本教科書

內容及國際試題，發展標準化成就測驗工具。題型先以選擇題為主，並視測驗目的，增加非選擇題部分，國語文部分則增加作文之施測。除發展各學科領域的測驗評量工具，亦將編製問卷，針對學生本人、家長、老師及學校行政人員進行問卷調查，以蒐集影響學生學習成就相關因素之資料。

(四) 實施方式

由本處結合國內學科、心理與教育測驗、統計之學者專家，先研發國小四年級、六年級、國中二年級、高中職二年級之國語(文)、英語(文)、數學、自然科學、社會等五個學科領域之標準化成就測驗工具，再視政策需要研發其他年級或學科領域之測驗工具。為儘早瞭解國小學生在國英數三科之學習成就，先於2005年4月底前完成小六學生國(含作文)英數三科之標準化成就測驗工具，並於2005年5月底正式施測。施測完畢後將試題公布在本處網站，學生可自行上網進行線上測驗。2005年5月底(含作文)對小六國英數抽測，2006年6月初針對小四、小六、國二、高中職二年級學生之國、英、數、自、社等五個學科領域進行施測。追蹤小組研究(panel study)則僅針對第一次(2006年6月初)抽測之小六學生，再抽取其中的10%學生，每兩年進行追蹤施測一次，直到高中職學校畢業為止(見圖一)。

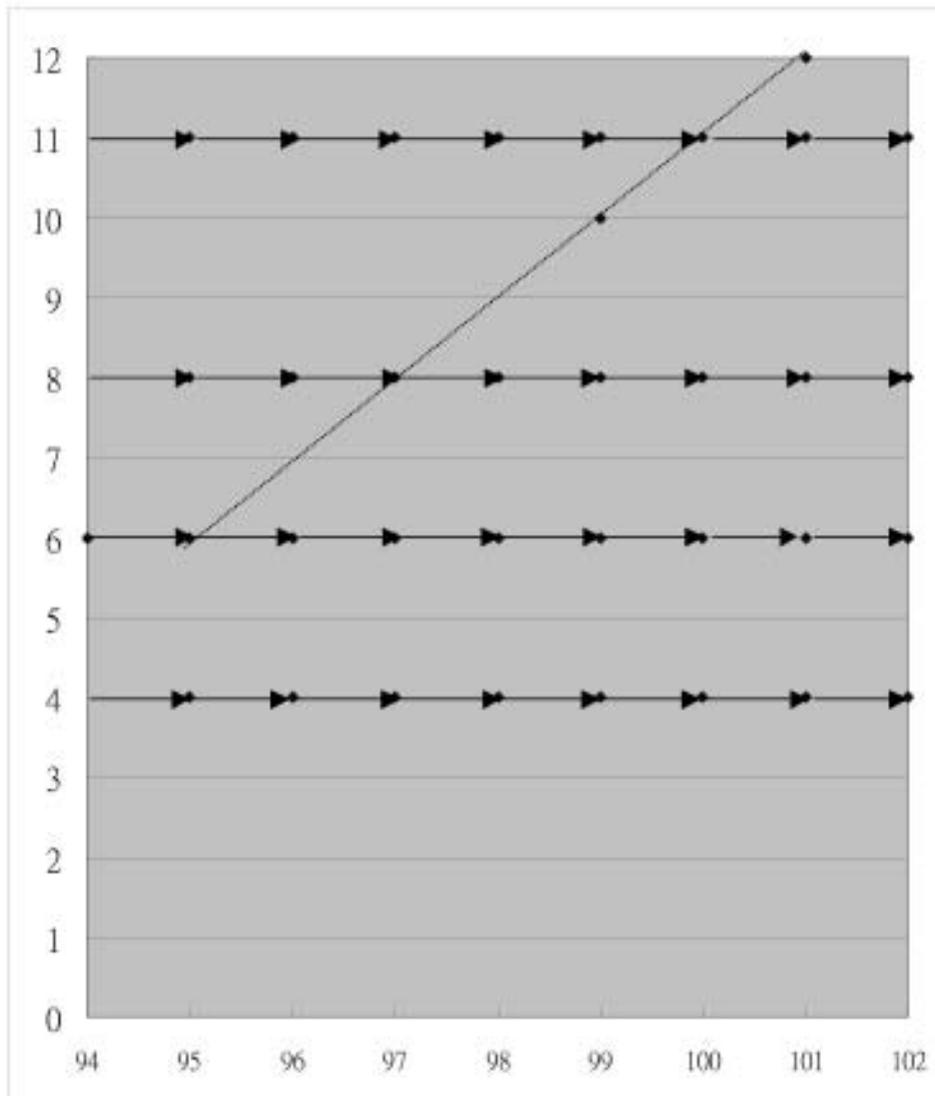
三、本資料庫之預期貢獻

本研究經由長期、整體的研究後，資料庫將可顯現下列的效益：(一)提供可靠且

便利的量化資料，瞭解國內教育狀況、學生學習成就變遷、品質與缺失，作為教育當局決策之依據。(二)提供豐富且完整的資料，供研究學者及學術機構研究分析。這不僅可大幅避免重複量化資料收集、降低資料收集彙整之花費與人力負擔，另一方面又可持有其統整性與代表性，進而呈現國內研究團隊之整合形象。(三)提供各教育階段進一步的資料分析，作為瞭解學生學習成效及學習狀況之參據。亦可作為實施補救教學之參考。(四)提供具有全國性代表的成就資料，不僅可以掌握國內學生學習成就表現，亦可供作國際比較，以便瞭解本國學生在國際上的狀況。

陸、結語

長期以來，各級主管教育行政機關或教育學者專家，都希望能建立整體與長期的教育資料庫，以提供教育當局進行課程與教學等決策的參考，亦能研究學者進一步進行資料分析、整理、組合、詮釋，以深入探討各種教育議題。教育部杜正勝部長上任後提出教育部未來四年施政主軸架構，其中「建立學生能力檢測機制行動方案」，其內容包括一、成立能力檢測單位。二、研擬能力檢測指標。三、建立學習成就資料庫。教育部乃委請本處研擬計畫，並同意本處所函報之「台灣學生學習成就評量資料庫研究計畫」，研究期程從93年10月至102年，本處正結合國內學者專家、縣市教育局、基層教師之力



圖一、台灣學生學習成就評量資料庫時序架構圖

註：◆ 指施測年

→ 指趨勢研究 (trend study)：每年進行小四、小六、國二、高二學生之抽測。

—◆— 指追蹤小組研究 (panel study)

X 軸：代表民國年

Y 軸：代表年級，國小一~六年級 (1-6)、國中一~三年級 (7-9)、高中一~三年級 (10-12)

