

與扶助，而在國內相關的方案中，雖然有如教育優先區或夜光天使專案等方案之推動，但當是以現行教育部積極著手進行的國民中小學之「攜手計畫」(<http://asap.moe.gov.tw>)，最具整合性、廣泛性與持續性。該計畫主要規劃對於學習落後之學生實施補救教學，使弱勢及學習成就不佳之學生，有提升學習成就之機會。而該計畫的實施，也成為國內國民中小學補救教學之實施依據，所以對於提升學生學習成就扮演著重要角色。

行政院於 2005 年核定教育部所提出之「攜手計畫」，該計畫主要目的為提供弱勢族群學生課業輔導，以縮小國民中小學學習成就低落學生之學習落差，並提供學習成就低落國中學生適性分組學習，以及多元學習方案。在該計畫方案中，已詳細規定各類實施要點，包含受輔對象的篩選方式、教學人員的資格、補救教學課程的內容與實施方式、考核機制等(教育部攜手計畫課後扶助網站，2009)。而為調查該方案是否有需精進之處，本研究曾另以問卷調查方式，針對國中小教師就攜手計畫之執行內容進行探討，探討項目包含「受輔對象」、「新增受輔對象篩選的方式」、「執行補救教學人選」、「配套措施」、「補救教學課程的內容適當與否」、「是否已達到預期提昇成就低落學生學力的成效」、「攜手計畫是否需要改進」等。最後，得總有效國中小教師樣本人數為 1933 名(詳細抽樣過程請見研究方法中之問卷調查)，而其分析結果，可提供中央主管機關作為改進攜手計畫之參考，在此僅簡要說明，以作為背景訊息，詳細可參考陳伯璋、林世華、謝進昌、陳清溪、曾建銘、林宜臻、蔡明學、謝佩蓉、周慧玲、金冠宇(2009a)之研究；研究中可以發現受訪教師認為受輔對象的篩選，應以客觀的測驗成績為依據，並加入教師對於學生的課堂表現進行判斷，而執行補救教學的教師應以受過教育專業訓練與實際經驗者為主，如果是現職教師，在配套措施方面受訪者認為，希望透過「降低原授課班級人數」、「降低授課時數」，以及「提供合理津貼」等相關方式進行。而補救教學課程內容可以由各校自訂，這樣較具教學彈性。不過受訪者同時也希望中央可以提供教材，或是訂立課程標準。最後在成效評估上，也希望中央能訂立一套評估標準，使學校與教師對於補救教學有明確的方向與指引。

### 三、國際現行學力品管機制

在評閱國際現行學力品管機制方面，本研究茲以美國、英國、日本與中國之品管架構為標的，探討如下：

#### (一)美國沒有落後的學生法案

在美國，教育原本是地方權責，但在中小學的部分，則有(美國)聯邦政府在 2001 年通過的沒有落後的學生法案(NCLB)法。NCLB 法案的立法，顯示(美國)聯邦政府對中小學的教育介入較深。其中，楊巧玲(2007)整理了美國教育的核心價值，由「公平」到「卓越」到現今 NCLB 法案的「標準本位」。各州政府各自擬定有一套因應 NCLB 政策中績效任務系統之內涵與執行的政策(例如張奕華

(2004)整理之 NCLB 下政府的密蘇里州的反應機制)或執行高中會考。大學教育幾乎是各校或各州(公立大學院校)自主，所以評鑑很少是全國性，許多評鑑措施，主要由學校自行決定，只有公立大學才會受州政府之一些規範與約束(彭森明，2008)。

NCLB 的緣由，係因(美國)國家教育統計中心(NEAP)的資料指出：非裔與拉丁裔十七歲的學生，其閱讀與數學的水平，僅與十三歲白人學生相等。因而，NCLB 法的目標，是預計在 2014 年前將此「學業成績的差距」加以接合。其中，NCLB 訂明：

- 1、每家學校必須到 2014 年，提升每位學生之數學和閱讀水平至其就讀年級水平(熟練)。
- 2、每家學校必須展示其在為每個子群組之目標，每年有足夠的進展(AYP)—這包括五個最大的族裔組，英語學習者，傷殘人士，和低收入學生。
- 3、接受聯邦政府 Title One 資助之高貧學校，如果在兩年內沒有達到 AYP，則必須支付補習費用或提供交通以便如學生想轉往別的學校。
- 4、同樣的學校如果學校在六年內沒有達到 AYP 要求，則必須重新組織，教職員換新，或作出一些大型之改變。

如今有關學生學力表現的相關統計報表，也公佈在 National Assessment of Educational Progress(NAEP)的網頁(<http://nces.ed.gov/nationsreportcard/nde/>)中，簡潔而有力的讓執政者與教育者很快的瞭解其中的訊息，例如在 NAEP 2007 年 Digest of Education Statistic 中，發表出兩個有關閱讀與數學公開的教育統計報表，隱約可從 1999 年的數據中，看出為何 2001 年會通過 NCLB 法案訴求的重要統計現象。以取自 NAEP2004 的資訊中，在 1999 年的數據中，可發現 17 歲的黑人在閱讀方面的平均表現是為 264(標準差 1.7)，而 17 歲的西班牙裔兒童的平均表現則為 271(標準差 3.9)，雙雙都只達 13 歲白人兒童的水準平均值為 267 分(標準差 1.2)；而在數學表現中，亦有類似的現象，顯示出非裔與拉丁裔十七歲的學生，其閱讀與數學的水平，只達到十三歲白人學生的學習成就，落後於同年齡的白人學生(NAEP, 2004)。在此背景之下，故到了 2001 年，NCLB 法案極欲想解決這種現象。然而，反觀 2004 年的數據，此種現象還沒有顯著性的改善，也許 NCLB 法案到 2004 時的實行的效應還沒有產生，如此，預期到 2014 年的報表中，能藉由這樣的統計資料，便可以明確的指出，NCLB 的政策下產生何種成效及是否達成其預定的目標。

美國在教育普及化之後，所有學童都可上學，由於地區與家庭環境差異等因素，造成學生個別差異，而各地學校教學標準與要求不一，因此學生學習成就參差不齊，若沒有統一標準，程度很低的學生也可以畢業。因此為了確保所有學生達到某種程度與標準才能畢業，各州除了要求學校教師加強教學與輔導，要求學生通過各科考試外，還可能要求學生通過全州統一會考。聯邦政府推動的 NCLB

法案也鼓勵各州推行高中畢業資格考試(High School Exit Exam, Graduation Exit Examination,..)，達到標準的學生才能畢業。目前美國約有半數的州有此規定(Department of Education and Early Development, Arkansas 2008)。畢業資格考試項目各科不大一致。以路易斯安那州(Louisiana)為例，其畢業資格考試(Louisiana Graduation Exit Examination)包括五科目：(1)閱讀與寫作知能(language arts)(2)作文(written composition)(3)數學(mathematics)(4)科學(science)(5)社會科學(social studies)( State of Louisiana,2008)。但加州只要求考閱讀與寫作知能(English language arts)數學(mathematics)和初等代數(algebra I)。

MCAS，1993年麻州通過「教育改革法(Education Reform Act)」，規定高中10年級學生必須參加「麻州綜合測驗(MCAS, Massachusetts Comprehensive Assessment System)」的英文與數學兩科，高中生除必須符合各學校區的畢業規定之外，尚必須通過這兩科會考之標準才可以獲得高中文憑(High School Diploma)，學生若未通過這兩科會考之標準，仍可以畢業，但不能獲得高中文憑。

麻州教育綜合測驗(MCAS)以實行十餘年，期間即使有不同意見，例如：某些教師反應增加教學負擔，有些教育學者批評考試干擾教學及有些能力不是考試可以測出，但麻州教育廳基於大多數家長的支持及為融合聯邦政府 No Child Left Behind 的政策，因此維持這項制度，並準備在2010年以後，高中生要獲得高中文憑必須加考自然科中的生物化學、物理，或工程科技中之任何一科。

「麻州綜合測驗(MCAS)」主要是根據麻州課程大綱(Curriculum Framework)的學習標準(Learning Standard)來評量學生的表現。評量結果分學生個人、學校、學校區三個層級作報告分析。麻州教育當局認為，MCAS可以有以下用途和功能：

- 1、改善教學如追蹤學生學習進展，如蒐集診斷性資訊、找出需要額外協助或補救教學的學生、發掘出目前課程和教學的優缺點。
- 2、學校和學校區績效責任：教育委員會(Board of Education)可以訂出改善的標準來看各校及各校區是否持續改善其教學成效。
- 3、學生的績效責任：學生提早被告知要通過英文和數學會考才可拿到高中文憑，又有重考之機會，因此此一會考也有激勵學生自我負責的功用 (駐波士頓文化組，2008)。

因應美國聯邦政府 NCLB 法案，各州政府皆擬定有一套 NCLB 政策中績效任務系統之內涵與執行的政策，也紛紛建立屬於各州的標準化測驗，以衡量學校的績效。由於 NEAP 數學與閱讀約隔四年會測一次，可監測全國及各州的成績與三個年級的表現，然而並不敷各州的要求，因而，各州紛紛開始建立標準化測驗，自增加測驗頻率，以至每年測驗並增加測驗的年級。此外，亦有某些學校採重視閱讀與數學的策略而排擠到社會領域的時數。

NCLB 法案要求每家學校必須展示其在為每個子群組之目標，每年有足夠的進展(AYP)，而這包括五個最大的族裔組，英語學習者，傷殘人士，和低收入學生。此從 NAEP2004 的統計數據可以看出，非裔與拉丁裔的平均分數，在低年級

是有進步的。然而，這個要求易使位於這類族群豐富的學區學校，很容易被不公平的評為不好的學校。

有關接受聯邦政府 Title One 資助之高貧學校，如在兩年內沒有達到 AYP，必須支付補習費用或提供交通如學生想轉往別的學校。這個部份 NCLB 實行以來，已有部份學校接受到這項處分的實施，然而移校就讀的效果如何，尚還有待追蹤研究的評估。至於同樣的學校如果他們在六年內沒有達到 AYP 要求，必須重新組織，教職員換新，或作出一些大型之改變。這項因為實行年限的關係還在觀察中。但有意見顯示應該把學校選擇與補充服務，應針對不能達到 AYP 的特定團體學生，而非把整個學校標籤為「失敗」。

美國的法案有充分自我監督機制，可以對下一次法案的重新授權作準備，但此部份在臺灣是相當薄弱甚至缺乏的(楊巧玲,2007)。本研究期望藉由美國 NCLB 以說明，一般需要追蹤或監控的統計數據，通常也是政策所關切的焦點。美國可以借鏡的地方是其標準化測驗的精神與長期追蹤的教育統計資料。而臺灣也可以確認我們想要的教育核心精神與重視哪些議題(縣市或特殊族群等..)並加以追蹤。臺灣自 2005 年開始建置臺灣學生學習成就資料庫 TASA，係屬於標準化測驗的精神，目前已在建置的階段，期許將來亦可建立具有貫穿各年級間的一把量尺教育報表，以供執政者與教育者參考。

## (二)英國評量架構

### 1、英國教育現況、目的與起源

英國在 1944 年的〈教育法〉中確立了教育的三個階段：初等教育(5~11 歲)、中等教育(11~15/16 歲)、繼續教育(16 歲之後，包括高等教育)，其中，中學教育階段屬於免費義務教育，學生可根據其年齡、程度及性向進入文法學校(grammar schools)、現代中學(secondary modern schools)、或技術學校(technical schools)等就讀(周愚文, 2008, p.51)。而根據英國政府的官方統計，2007 年 1 月英格蘭小學班級的平均生師比是為 26.2:1；在教學上，小學通常為不分組教學，沒有能力分班，僅少數科目或部份教學方法採班級內分組進行的方式。

在課程方面，是於 1988 的〈教育改革法案〉開始採用「國定課程」(National Curriculum)、並首次定義了義務教育中的「關鍵階段」(key stages)，其下主要分為 4 個階段，而後國定課程評量(National curriculum assessment)即依各個階段，設立能力標準，建立起各年級評量的基礎(DCSF, 2008)。

在權則分配方面，由於英國是屬地方分權，評量的方向與原則是由中央訂立，但實際作法與細節(如教師評量)則由各學校自行決定，搭配政府與家長的監督以建立起國家的評量網絡。

### 2、評量架構

英國的評量體系，主要是在義務教育的時程中，分成 4 個主要的關鍵階段，如表 4-9 所示，而每個階段下都有其重點評量科目與評量方式，其中評量科目是

由國定課程中所組成，而評量方式主要區分教師評量與國家測驗兩種。以下茲分初等教育與中等教育分述之：

表 4-9 英國各關鍵階段之評量方式一覽表

年齡	階段	年	評量方式
3-4	學前階段		
4-5			
5-6	Key Stage 1	Year 1	
6-7		Year 2	英語、數學、科學的教師評量
7-8		Year 3	
8-9	Key Stage 2	Year 4	
9-10		Year 5	
10-11		Year 6	英語、數學、科學的國家測驗與教師評量
11-12		Year 7	
12-13	Key Stage 3	Year 8	
13-14		Year 9	英語、數學、科學的國家測驗；其它基礎科目的教師評量
14-15		Year 10	某些學生接受 GCSEs
15-16	Key Stage 4	Year 11	大部份學生會接受 GCSEs 或其它國家資格檢定考試

註：GCSEs：中等教育認證科目(General Certificate of Secondary Education)

#### (1)初等教育

一般來說，英格蘭的小學招收 5~11 歲的學生，如表 4-9 所示，是屬於英國國定課程中的第一關鍵階段 (key stage 1, 5~7 歲) 和第二關鍵階段 (key stage 2, 7~11 歲)。英格蘭的初等教育皆採用英國國定課程。根據國定課程的第一關鍵階段及第二關鍵階段，英格蘭初等教育的核心科目包括英語、數學、和科學，此外還有基礎科目包括：資通科技 (information and communication technology, ICT)、設計與科技 (design and technology)、歷史、地理、藝術與設計 (art and design)、音樂、體育 (PE)、以及宗教教育。

除國定課程之外，英格蘭的小學還規劃許多非國定科目或課程，例如經濟、環境教育、政治與國際 (包含歐洲認同、歐洲歷史文化與社會經濟)、公民與品德教育、性教育、藥物教育等。此外，英格蘭已規劃在 2010 年之前，全面提供第二關鍵階段的學生現代外語教育。而威爾斯的非國定科目或課程則是將重心擺在對其文化、經濟、環境、歷史、及語言等特性的認識和了解。

在學習評量方面，架構如表 4-9 所示，在每一關鍵階段結束前，都會針對核心科目進行評量，而這些科目包括：語文（英語）、數學、和科學。其中，評量方式，英格蘭小學在第一關鍵階段結束前（7 歲），是由教師對學生的學習情形對照國定課程能力指標（attainment targets）的設定進行等級評量，而對於等級評量，如表 4-10 所示，是將整個義務教育分成 8 個層次（詳細如附錄一所示），其中，大部分初等教育學生的等級是介在第一級到第三級之間，而初等教育的目標乃希望學生在第一關鍵階段結束前達到第二等級。而第二關鍵階段結束前（11 歲）則是同時利用成就評量（即國家測驗）及教師評量來評定等級，其中，大部分學生的等級多落在第二級到第五級之間，而教學的目標則是希望學生能達到第四級，此外，若是學生無法達到該預期等級，則必須進一步接受補教教學。

表 4-10 學生於各關鍵階段預期達到的水準

關鍵階段	水準範圍	年齡	預期達到的水準
Key stage 1	1-3	7	2
Key stage 2	2-5	11	4
Key stage 3	3-7	14	5/6

## (2) 中等教育

英格蘭的中等教育範圍係指提供給 11~19 歲學生的各類型學校教育，其中 11~16 歲為義務教育。中等教育的前五年（11~16 歲）屬於義務教育，其中 11~14 歲是屬國定課程的第三關鍵階段，而 14~16 歲是為第四關鍵階段。

英格蘭的中等義務教育包括國定課程、宗教教育、以及各校自訂之課程，且所有接受公款補助之中等學校（maintained schools）皆應提供性教育課程，內容應至少包括對愛滋病、愛滋病毒（HIV）、和其他性傳染病的認識。此外，中學也必須提供毒品菸酒教育、工作相關的學習、和生涯教育。

英格蘭第三關鍵階段的國定課程包括核心科目（英語、數學、和科學）和基礎科目（設計與科技、資通科技、歷史、地理、現代外國語（modern foreign language）、藝術與設計、音樂、體育、公民）。國定課程之外，自 2000 年 8 月起，中學生都應修習品德與健康教育（personal, social and health education, PSHE）；而公立學校都應提供宗教教育課程，但家長有權力決定自己的孩子是否參加此課程。而在評量上，如表 4-9、4-10 所示，是由國家測驗與教師評量來評定學生在整個能力層級中，是否達到預期的第 5/6 的水平。

到了第四關鍵階段，學生可以選擇自己想要修習的中等教育認證科目（General Certificate of Secondary Education, GCSEs）（Directgov, 2008a），可選擇的範圍包括超過 40 種一般科目，以及 8 種應用類科（與工作相關的學習科目），這些科目有部份與國定課程重疊。一般而言經過五個學期（將近兩年，也就是 14~16 歲）的學習後便可參加中等教育認證考試，各校也會視學生的程度而允許資賦優異的學生提早考某些科目。中等教育認證考試不限定每次應考幾科，許多

科目的成績會包括平時學習成果及考試結果，成績評定以八個等級表示：A\*、A、B、C、D、E、F、G，若成績不達最低標準則標示為 U，表示無法拿到認證。

根據 2002 年教育法案，英格蘭的第四關鍵階段國定課程開始與前三階段有明顯區別，使該階段的義務教育更具有彈性，並自 2004 年九月起頒佈實施，其中的改變包括：(1) 減少義務教育的核心科目，目前的核心科目有：英文、資通科技、數學、科學、公民、體育、和宗教教育。(2) 將生涯教育與性教育囊括進義務教育的學習領域。(3) 所有接受公款補助之中等學校皆應提供與工作相關的學習。(4) 除了核心科目之外，各校應就以下四領域：藝術、設計與科技、人文學科、及現代外國語，各領域至少應提供一門課程。此一「14~16 歲學生增加彈性方案」(Increased Flexibility for 14- to 16-Year-Olds Programme, IFP) 爲了使學生有更多職業或工作相關學習的機會，還請當地的擴充教育學院與這些中學及代辦單位合作，使中學生有機會到學院上課，或者每週安排一至兩天接受職業訓練教育，學生也因此能爲職業或工作相關的證照考試作準備和學習，並發展其對工作／職業狀況的認識和了解。

16 歲~18 歲以上是屬義務後教育，學生所接受的義務後教育不在國定課程的範圍，因此學生學習的課程內容將依照其所選擇的認證考試及認證單位所公佈的課程大綱來安排，可能的選擇包括：進階級子認證 (Advanced Subsidiary qualifications, GCE AS)、進階級教育認證 (GCE A-level)、進階級職業教育認證 (A-levels in applied subjects)，總共包括了超過 80 種學術科目及職業科目 (Directgov, 2008b)。

### (三)日本評量架構

#### 1、日本全國學力調查沿革

日本於 1960 年代就已展開全國性學力測驗，由於學校及地區競爭過熱，1964 年停止全體調查，1965 年起改以抽樣方式，由於國家層級舉辦之學力調查被判定違法，1966 年起全面停止。1982 年起以局部之中小學生爲對象，2002 年起以局部之高中生爲對象，再度展開學力調查。2004 年當時的文部科學省(相當於我國教育部)中山成彬部長爲因應學力下降以及確保次世代人材基盤的義務教育品質的提升，客觀掌握學童學習狀況的必要性考量下，2005 年 6 月「2005 經濟財政營運和改革構造的基本方針」的內閣會議(首相官邸，2005)以及中央教育審議會答詢「創造新時代的義務教育」決定實施全國性學力調查(文部科學省，2005a)，同年 11 月文部科學省召開全國性學力調查實施方法等之專家檢討會議(文部科學省，2005b)，討論具體實施的方法(文部科學省，2006)。睽違 43 年，2007 年日本再度以中三及小六全體爲對象(※僅少數學校以引發競爭爲由拒絕參加)，展開國語與數學的學力調查 (文部省，2007)，以掌握分析各地區學童的學力與學習狀況，檢討教育及教育施策之成果與課題以求改善，提升學童之學習效果及提升學習意願，進而保障義務教育機會均等及維持與提升全國水準。

## 2、以普查方式實施之緣由

學力低落問題受到重視下，基於學校教育的現狀和課題充分地把握的必要性、國際學力調查結果發現學力和學習意願有下降傾向、建立保證義務教育質量架構的請求，而實施學力調查。學力調查採取全國普查方式實施的原因在於：全國學力狀況趨勢分析的抽樣性調查，對各學校教學助益不大，而全國性普查除提供教育政策的確立與「學習指導要領(指日本的課程標準)」修訂等之依據外，可以提供地方教育委員會、學校以及教師改進教學；其次，國際評比排名與學習意願下滑，例如：日本過去在 TIMSS 與 PISA 等國際評比的數學、自然科學等成就測驗上，有著良好的成績表現，日本學生的 TIMSS 小四數學排名：1995 年與 2003 年第 3 名、2007 年第 4 名；八年級數學 TIMSS 1995 第 3 名、TIMSS-R、2003 與 2007 都是第 5 名。日本 PISA 數學排名則從 2000 年的第 1→2003 年的第 6 下降至 2006 年的第 10 名(林宜臻，2009)；最後，建立保證義務教育質量架構的請求：1980 年代以來，日本為改變填鴨教育實施「寬鬆(裕)教育」政策，1998 年版的學習指導要領將原有教學內容削減了三分之一，該政策被認為是導致日本學生學力下降的重要原因，因此，為提高學校和學生的競爭能力及促進學習，文部科學省實施全國性學力普查。

## 3、調查目的

- (1)全國義務教育機會的均等與水準的提升，由各地區學生的學力與學習的掌握與分析，檢驗教育與教育施策的成果與課題，以謀求改善。
- (2)各教育委員會、學校等從與全國狀況的關係中，掌握各自教育及教育施策的成果與課題，謀求改善的同時，確立後續的檢證與改善的機制。
- (3)各校得以掌握學生的學力與學習狀況，幫助對學生的教學指導和學習狀況的改善等。

## 4、調查內容

其調查內容，主要以學科中的國語、算術/數學為主。

(1)出題範圍：原則上以該調查年級先前的學習內容為主。

(2)出題內容：以「知識」和「活用」為出題種類。

「知識」算術/數學 A 以會影響後續學年學習的內容及生活中活用不可或缺的知識、技能為主。

「活用」算術/數學 B 將知識、技能等活用於生活中各種場面的能力，以及為解決各種課題的構思及評鑑、改善能力等之內容為主。

問卷調查(生活習慣和學習環境等)其下分為學生與學校問卷，其中，學生問卷：以接受調查該學年的學生為對象，調查其學習意願、學習方法、學習環境及生活等；學校問卷：調查學校的指導方法以及學校的人力、物力上等教育條件之整備狀況。



## 5、實施方式

本調查以紙筆測驗及問卷調查方式進行，以瞭解學生對語文及數學的基礎知識與基本技能的掌握和活用情況，以及相關的背景和條件，本文將就下表 4-11 之 2009 年 4 月 21 日進行的學力調查為例探討之。

國立學校無論小六或九年級階段全數參加，公立學校也幾乎全數參加，惟 2009 年九年級階段下降至 96.9%，私立學校無論小六或九年級階段參加比例最低，由六成下降至 43.1%(小六)及 46.6%(九年級)。

表 4-11 日本學力調查範圍

校別	年度		
	2007 年	2008 年	2009 年
國立學校	75 校 (100.0%)	76 校(100.0%)	76 校(100%)
小六 公立學校	21,889 校 (99.8%)	21,685 校(99.8%)	21,482 校(99.8%)
私立學校	108 校 (60.0%)	88 校(47.1%)	85 校(43.1%)
合計	22,072 校 (99.5%)	21,849 校(99.4%)	21,643 校(99.3%)
國立學校	81 校(96.3%)	79 校(97.5%)	77 校(95.1%)
九年級 公立學校	10,050 校(98.0%)	10,121 校(99.3%)	9,851 校(96.9%)
私立學校	416 校(60.5%)	374 校(53.4%)	330 校(46.6%)
合計	10,544 校(95.7%)	10,574 校(96.4%)	10,258 校(93.6%)

## 6、日本學力調查的特色

### (1)精進調查工具

#### a. 試題內容生活化

日本借鏡 PISA 國際學力調查，將試題除「知識」外，同時將「活用」列入試題是一大特色，A 卷的「知識」主要考查學生不學會將影響後續學年的學習內容，以及生活中活用時不可缺少的知識與技能為主。B 卷的「活用」主要考查學生能否將知識與技能等活用於實際生活中的能力以及解決課題的構思、評鑑、改進等的的能力。

#### b. 題型多樣

A 卷由選擇題、簡答題組成，B 卷由選擇題、簡答題和問答題組成。

#### c. 問卷內容廣泛

問卷調查包括對學生和學校相關背景的調查，除了涉及學生的學習興趣、態度、學習方法等，還有學習環境、生活等方面的內容。因此，調查結果不僅可用

於學校教育，也有利於學生家長協助學校教育，改善學生的生活習慣、家庭環境等影響學力的因素。

#### (2)變更調查年級

爲了實施全國學力調查，文部科學省於 2005 年 11 月 8 日設置「全國性學力調查實施方法等之專家檢討會議」(文部科學省，2005b)歷經七次會議討論具體實施的方法，翌年(2006年)4 月公佈了關於全國學力調查具體實施方法的報告(文部科學省，2006)。較以往之相關調查係針對小學五年級學生進行，該報告指出，爲了掌握義務教育水準，選擇小學六年級及九年級較爲合適，由於調查在新學期 4 月進行，因此測驗考查範圍原則上以調查年級之前的指導事項。

#### (3)調查結果同年公佈

學力調查於 2009 年 4 月 21 日進行，2009 年 8 月 27 日就公佈調查結果，12 月提出調查結果報告書。

#### (4)公開調查問題

日本教育部於調查實施後，除提供統計分析之相關數據外，亦公開調查問題、正確回答事例、題目主旨以及解答類型。

#### (5)規範公開方式

日本會公佈全國性調查結果，包含公佈全國、都道府縣、地區規模(大都市 核心城市、其他城市、鎮村、邊遠地區) 的調查結果，其次，分析與公佈學生的生活習慣及學習環境、學校教育條件的整備狀況等與學力之相關情形。

#### (6)提供給教育委員會、學校等調查結果

考畢，日本會提供都道府縣、市鎮村、學校等之調查結果，同時，不歸還學生答案卡，但透過學校提供個人正確回答與誤答狀況之通知，此外，不明確指出各個市鎮村名、學校名等，避免學校排名及過度競爭。但市鎮村、學校得以向家長等說明自己的結果，最後，調查結果不作為入學考試資料用。

#### (7)提供教學改進建議

評量結果除指出學童答對率外，並分別針對內容指出學習上之問題點，並提出教學改進之建議。

### 7、日本學力調查的問題點

#### (1)「活用」型的問題仍待努力

2009 年公佈的學力調查結果發現，秋田、福井、富山等縣市因年仍名列前茅，表現好的縣市，除了學校指導有力外，家庭中的學習習慣及正確的生活習慣，

也是影響學力的提升。原先排名最差的沖繩及大阪等縣市，向排名第 1 的秋田縣學習，沖繩甚至與秋田交換教師，大阪市加強計算問題的反覆練習，小學階段屬於「知識」類型的問題，成績接近平均值，排名也由 41 名提升至 33 名，但「活用」的問題仍待努力，中學則連續三年全國最後(產經新聞，2009 年 8 月 28 日)，

#### (2)鄉鎮村別及學校別的資訊是否公開

規制改革會議 2008 年 12 月的答詢中指出「使用多數的公費未提供國民相對的資訊」，內閣府委託民間調查公司於 2009 年 1 月進行調查，結果發現：約 7 成的家長認為應以學校別公開成績，約 9 成的縣市教育局則主張不宜公開(朝日新聞，2009 年 8 月 3 日)。鳥取縣教育委員會無視於文部科學省要求不能公開學校別及市町村別(指鄉鎮村)的調查結果，首開公佈學校別及市町村別全國學力調查結果之例(読売新聞，2009 年 9 月 08 日)。

#### (3)全國性普查是否合宜

全國性普查的方式 2007 年耗資約 77 億日幣，2008 年及 2009 年耗資約 58 億日幣，2010 年的預算也達 57 億日幣，被指出「分析所得結果與經費和勞力間不成正比」(朝日新聞，2009 年 8 月 24 日)，民主黨建議以抽樣調查方式取代全員參加的學力調查(產經新聞，2009 年 9 月 2 日)，隨著 2009 年 9 月 16 日民主黨鳩山由紀夫新政權的誕生，教育部長川端達夫指出重新檢討以抽樣方式進行調查以及增加科目等(每日新聞，2009 年 9 月 17 日)。

#### (4)體制的矛盾

梅原利夫(2007) 指出(1)測驗分數的高低成爲唯一的評量基準—教育課程由各學科、道德及特別活動構成，片面性的評量基準違反教育宗旨；(2)與「多樣學力與生存力」的學力觀矛盾—爲提高 PISA 調查成績。(3)委由民間公司參與獨攬造成不安—評量題編制雖由文部科學省負責，但試題及答案用紙的印刷、梱包、發送、回收、計分、和數據處理、公告資料製作等作業則委託精通考試的民間教育產業擔負，其中結果的數據及問卷的個人情資由民間公司獨佔引發不安。(4) 外部評鑑扭曲教育評鑑—評鑑應由自己評鑑、相互評鑑、外部評鑑、綜合評鑑等機能性組合之，方利於教育活動的有效進行，由外部評鑑取代評鑑，教師與學校目標由外部設定，將不能因應事實進行適切的診斷。

### (四)中國評量架構

「建立中小學生學業品質分析、回饋與指導系統」專案與「國家基礎教育品質監測中心」共同構成了中國國家基礎教育品質保障系統。

#### 1、中小學生學業品質分析、回饋與指導系統

##### (1)建立的背景

中國隨著基礎教育已從普及邁向教學品質全面提升的新階段，而教學品質大範圍的提高，有賴於科學的宏觀調控和指導，在基礎教育課程改革深入推進的階段，適應新時期基礎教育發展的需求，2003年，在教育部基礎教育司的指導下，中國教育部基礎教育課程教材發展中心組織成立了「建立中小學生學業品質分析、回饋與指導系統」專案組。其成員包括課程教學、心理學、教育測量、教育統計、電腦軟體程式設計等各方面的專業人員。中國教育部課程中心主任朱慕菊針對中國當前沒有一個基礎教育的資料平臺，不能對各地區的基礎教育情況作橫向的比較的現狀，力圖通過建立國家級、省級常模，進行歸因分析，指導教學，從而促進義務教育更均衡地發展，指出「建立國家中小學學習品質分析與指導系統」的重要性(中小學學習品質評價的新動向，2007年3月8日)。

## (2)目的

「建立中小學生學業品質分析、回饋與指導系統」專案設置目的在於：(1) 真實反映學生學業品質狀況；(2) 提高教育決策科學化水準；(3) 教育教學診斷和品質提高提供依據；(4) 提升教師和相關人員的專業水準；(5) 引導社會樹立正確的品質觀。

對學生學習品質進行評量監測，主要是為決策和指導服務，其意義在於：(1) 反應學科教學多方面的指標資訊；(2) 對學校和地區教學過程進行診斷分析；(3) 測試之後對學校和地區教學工作進行針對性指導(董洪亮，2008年2月18日)。針對中小學生學業品質進行檢測、分析、指導。經由資料的收集、分析和回饋，幫助各級教育行政、教研部門，有針對性地改進教學中的問題，提高學業品質。來自一線的優秀教師的教研人員，面對全面提高基礎教育品質，在經驗積累的基礎上，學會運用科學的手段。

## (3)實施概要

中國的評量架構，其調查對象是以小學三年級和中學八年級，而其調查內容包含學科調查與問卷調查，其中，學科調查是以三年級測試語文、數學兩個學科、八年級測試語文、數學、英語、科學四個學科，而問卷調查是包含：學生問卷，以調查的內容包括家庭環境、學校環境、課業負擔、師生關係、與同學的關係、學習方法、學習態度等；教師問卷，調查的內容包括教師基本情況如學歷、任職經歷、職稱，教師對新課程的看法，教學策略和教學方法，教師對學校管理的看法，教師面臨的困難，教師的教育觀和教師專業發展等；最後，校長問卷，調查的內容包括學校硬體條件、校長辦學理念和教育觀、校長的學校管理方法等。

評量架構方面，各學科學習品質評量架構主要依據課程標準，考核學生對學科核心知識與技能的理解和掌握，尤其重視科學方法的掌握、解決問題的能力、收集與分析資訊的能力和反思批判能力。評量架構是將課程標準中對學生的要求轉化為操作化、行為化的描述，將之分為內容與能力兩個向度。

#### (4)特色

##### a.命題內容的特色

以課程標準為依據，在基礎知識與基本技能的雙基基礎上，重視核心素養與核心能力，使用真實的情境和任務，以客觀性試題考察學生高層次認知能力，突顯學生未來發展的核心素養和核心能力。其命題內容包括：(1) 學科核心知識、技能的理解和掌握；(2) 綜合運用所學知識，解決實際問題的能力；(3) 收集與分析資訊能力、反思批判的能力，以及對重要學科方法的掌握。

##### b.試題編製的特色

每一道試題編製的都經過專家審議和多種形式的試測，經過層層篩選，明確標定每一道題目的難度、鑑別度、反應曲線等資訊，並運用布魯姆分類法對能力進行界定，同時結合索羅分類法 (SOLO Taxonomy)設計題目選項，以測驗的方式瞭解學生解決問題時表現出的認知發展水準，更精細區分學生能力。

##### c.評量結果解釋的特色

評量結果的解釋同時結合標準參照與常模參照，標準參照確定學生處於怎樣的水準，常模參照比較不同類別學生的表現，使教育行政和教研部門等超越個人經驗的侷限性，從不同的層面，瞭解學生在學業上的水準和特點，能更準確地判斷自己地區或學校各學科的整體水準及優弱勢，得以有針對性地進行改進和指導，還可以利用資料資訊追蹤研究，連續若干年瞭解學生學業品質的發展變化狀況。

##### d.測驗結果處理的特色

###### (a)測試專家分析而不解說因果關係

專案組負責命題、抽樣、測試過程管理、閱卷、資料錄入與整理、資料分析、報告撰寫，以及教學改進等項目，呈現學生在各學科不同內容領域和能力維度上的表現及其在典型題目上的作答情況，分析學生學業成績和學習環境因素的關係，培訓一線教研員幫助解讀報告，過程中，測試專家只對數量關係進行分析，因果關係的解說由當地的管理者、研究者和教師承擔。

###### (b)即時回饋系統

12月底完成測試，隔年3月前後，完成資料處理和報告撰寫。4月，國家專案組組織專家對全省的測試情況進行省一級的回饋。5月，國家專案組支援所有測試科目教研人員的專門研修活動，進一步學習科學命題、科學評量的相關技術內容，更深入地解讀省、市、縣區三級的分析報告。6月底，省教研室組成專家組，進行巡迴式的回饋活動，並要求各地級市提交利用測試結果改進教學工作的規劃，規劃當中包括測試結果、原因分析、改革重點以及改革的計畫和保障等等內容。

### (c)依據測試結果展開課堂教學改革活動

各地教育行政部門以對測試結果的分析為依據，展開一系列針對課堂教學的改革活動。以江蘇省為例，在測試結果分析的基礎上，省教研室組成專家進行巡迴式回饋活動，同時要求各地級市提交利用測試結果改進教學工作的專門計畫和報告。在市級回饋的基礎上，組織了縣區級的回饋、研討與交流活動。實證資料面提供當地教育的優勢和不足，市教育局組織政府領導、教育行政部門領導、學校校長等為學科教學品質進行專門研究，利用 3 個月時間對所有小學的教學狀況進行拉網式調查研究，瞭解師資情況、設備條件情況和課堂教學情況，推動教學管理與課堂教學改革，針對性執行改進措施。

### (d)資料分析報告內容特色

提供的資料分析報告內容包括：學生在各學科上的總體表現；學生在各學科不同內容領域和能力層次上的表現；不同群體學生的比較；學生在每題的作答情況；典型題目分析；問卷調查結果分析。根據測試結果的統計分析，以學校或區域、學科為單位，針對學習狀況、學習能力、教學過程、管理過程進行實際診斷，並提出相關的改進建議。

### (e)由需求者自籌經費提出申請參加測試

基於「中小學學習品質分析與指導系統」專案裨益於：提高教育行政部門、教研部門的決策和指導水準；促進教學品質的大範圍提高和基礎教育的均衡發展；引導評量方式的轉變。以江蘇省為例，2006 年年底，江蘇省教育廳投入 600 萬人民幣左右的專項經費實施覆蓋全省 13 個地市、106 個區縣 11 萬多人次的抽樣測試(董洪亮，2008 年 2 月 18 日)。

### (f)利用定錨題探討趨勢

以課程標準的要求為參照，為保證歷次採集到的資料本身具有連續性和可比性，題型、題量以及每一類題型的檢測重點等方面具有高度的穩定性，而且每一次測試的試題都與「定錨題」的類型相同，瞭解學生學業品質的發展變化狀況。

## 2、中國教育部基礎教育品質監測中心

### (1)緣起與角色

相對於中國教育部基礎教育課程教材發展中心於 2003 年組織成立的「建立中小學生學業品質分析、回饋與指導系統」專案組，2007 年 9 月中國中央機構編制委員會辦公室批准成立「教育部基礎教育品質監測中心」，在教育部的直接領導下委託北京師範大學建立專業機構。中國胡錦濤總書記在 2007 年 10 月 15 日在共產黨第十七次全國代表大會上的報告重申「優先發展教育，建設人力資源強國」，強調實施素質教育，更新教育觀念，深化品質評量制度改革，建立與新的品質觀相適應的評量制度，以保證學校基礎教育品質的全面提升，指明教育督

導評估工作的發展方向 (鄭富芝, 2007 年 11 月 09 日)。同年(2007 年)11 月成立「教育部基礎教育品質監測指導委員會」, 由課程、學科和教育評量、學業測量等方面的專家組成, 為教育品質監測工作提供學術諮詢和支援, 用以加強對基礎教育品質監測工作的宏觀指導與管理, 確保教育部基礎教育品質監測中心權威性和科學性。中心負責擬定基礎教育品質監測標準, 組織實施全國基礎教育品質監測工作, 為各地展開基礎教育監測工作提供技術支援和業務指導。

## (2)目的

陳小嫻教育部副部長在由聯合國教科文組織國際教育規劃研究所與上海市教育科學研究院聯合主辦的「變遷世界中的成功教育體系：教育系統效能監測與評估」政策的研討會指出：基礎教育監測是科學決策的必然要求, 是進一步轉變政府職能、改善宏觀管理的重要措施(王軼軒, 2009)。中國教育部部長周濟在教育部基礎教育品質監測中心揭牌儀式上, 指出要建立基礎教育品質監測體系, 全面準確把握基礎教育的品質狀況, 科學診斷我國基礎教育品質存在的問題和原因, 教育決策者可以根據監測所提供的資訊對未來的教育進行調整。(佚名, 2008 年 2 月 14 日)。對基礎教育品質進行全面的監測和科學的評量, 全面準確掌握學生的真實素質狀況, 其目的在於：(1) 探索人才培養的規律, 促進各地區、各學校全面實施素質教育；(2) 提供實踐和理論依據, 全面提高教育品質；(3) 提高教育督導的針對性和時效性, 使教育督導工作監督有據、指導有方, 提高督導的服務水準。

## (3)實施方式

### a.監測內容

周濟教育部部長指出基礎教育品質監測要：(1) 按照基礎教育品質標準進行；(2) 監測範圍要覆蓋學生德、智、體、美全面發展的各個方面；(3) 要特別重視監測學生的創新意識和心理健康水準。監測的重點領域包括學生的身心健康水準、學生的綜合素質、學生的學業成績, 考查中小學生在品德、智力、體質等方面的全面發展狀況, 以及獨立思考能力、創新意識、實踐能力、影響學生發展的教育環境等(佚名, 2008 年 2 月 14 日)。

### b.監測對象

以抽樣方式對基礎教育階段所有在校學生進行整體監測。

## (4)特色

### a.探索教育發展的規律

「依法監測, 服務決策」原則下, 以法律/法規以及基礎教育課程改革標準為依據, 採用科學的監測方法, (1) 診斷基礎教育品質存在的問題和原因, 幫助尋求解決方案；(2) 探索基礎教育發展的規律、人才成長的規律和教育管理的規

律。

#### b.綜合監測

國家基礎教育品質監測中心主要從：(1) 學生的思想品德和公民素養；(2) 身體和心理健康水準；(3) 學業水準和學習素養；(4) 藝術素養；(5) 實踐能力和創新意識；(6) 影響學生發展的教育環境與社會環境等六個方面進行綜合監測。

#### c.提供支援和指導

中心負責擬定基礎教育品質監測標準，組織實施全國基礎教育品質監測工作，為各地展開基礎教育監測工作提供技術支援和業務指導。

### 3、中國評量架構所提供之經驗、建議

#### (1) 設置改善與監控體系

中國為保障國家基礎教育品質系統，設置「國家基礎教育品質監測中心」與「建立中小學生學業品質分析、回饋與指導系統」。「國家基礎教育品質監測中心」屬於外部的「品質評估與監控」系統，探索教育發展的規律，面向政府報告基礎教育品質的現狀，為教育決策提供資訊、依據和建議。「建立中小學生學業品質分析、回饋與指導系統」面向基層，為基層教育行政、教研部門直接提供所有資料資訊，整個過程對一線教師開放，是一種內部的「品質保障與改善」系統，著眼於基礎教育品質的改善和保障。為教育系統內部和外部改善與監控各執其司下，得以「問責」與「回饋」，保證國家基礎教育品質並促進其不斷提升。

#### (2) 執行綜合監測

監測機制一般著眼於學業水準和學習素養，中國的「國家基礎教育品質監測中心」則針對基礎教育階段的學生德、智、體、美以及影響學生全面發展的環境因素等進行綜合監測。

#### (3) 提供科學化測試系統

提供科學化測試系統，轉變政府職能、改善宏觀管理，使教育行政和教研部門等超越個人經驗的侷限性，瞭解當地教育的優勢和不足，以及學生在學業上的水準和特點。此外，由需求者自籌經費提出申請參加測試，以免各地自行架設測試系統，疊床架屋。

#### (4) 建立監回饋與指導系統

中國的「建立中小學生學業品質分析、回饋與指導系統」，進行教學診斷並指導教學，幫助地方與學校逐步提升品質，重視分析與指導，以保障課程計畫的落實和達到課程標準要求。



(5) 分析而不解說因果關係

只對數量關係進行分析，因果關係的解說由當地的管理者、研究者和教師進行解說，將更能因地制宜，進行適切的診斷，能更準確地判斷自己地區或學校各學科的整體水準及優弱勢，得以有針對性地進行改進和指導。

(6) 即時回饋

各地教育行政部門方能依據測試結果，即時展開針對性的對策，進行課堂教學改革活動。

(7) 提供學生學業品質分析報告

提供呈現學生在各學科上的總體表現及其在不同學科內容領域和能力維度上的優勢和不足，並分析了學生學業成績和學習環境因素的關係的學生學業品質分析報告，方利於課堂針對性教學。

(8) 試題內容生活化

日本學力調查，將試題除「知識」外，同時將「活用」列入試題，「活用」的試題主要考查學生能否將知識與技能等活用於實際生活中的能力以及解決課題的構思、評鑑、改進等的的能力。

(9) 檢測認知發展水準

結合索羅分類法 (SOLO Taxonomy) 設計題目選項，以測驗的方式瞭解學生解決問題時表現出的認知發展水準，更精細區分學生能力。

## 四、本建議方案之思路與國際趨勢

在前文，本研究已詳盡描述國內相關的品管方案(例如：攜手計畫)與美國、英國、日本等國家之學力品管架構，而面對如此多元與豐富的資訊，要如何有效的規劃或者從哪裡開始執行?則變成一個重要的課題。研究團隊在依循 Shavelson, McDonnell, Oakes, Carey 與 Picus(1987)的想法下，以他們所認為監控系統的資訊來源是包含下列五種，作為參考建議，更指引本研究建置學力品管之思路，作為未來執行的方向，陳述如下：

### (一) 採用現行可得資料(Status quo)

對於監控系統的資料來源，Shavelson, McDonnell, Oakes, Carey 與 Picus(1987)初步的想法很簡單，他們認為如果現行的資料已經能夠滿足或回答所需的問題時，研究者可直接採用，不必另行開發或建置，但這方面的資料，來源可能非常複雜(或時常變動)，必須視當下時間，所能夠獲得哪些訊息為主。

### (二) 機制的整合(Patchwork)

Shavelson, McDonnell, Oakes, Carey 與 Picus(1987)所提出的另一項選擇，是