2006年台灣學生學習成就 評量結果之分析

曾建铭

國家教育研究院籌備處助理研究員

陳清溪

國家教育研究院籌備處主任秘書

摘要

台灣學生學習成就評量資料庫(TASA)已邁入第三年,本研究主要在呈現國語文、英語文、數學、社會及自然等五科在國小四年級、國小六年級、國中二年級、高中二年級及高職二年級等五個教育階段在2006年全國分層隨機抽樣施測的結果。施測的結果包括九年一貫能力指標與高中、高職課程綱要內容的答對率描述性統計分析,以及學生相關背景變項的統計分析。其結果除了提供學生、教師與家長參考外,最重要的是提供具體的學生學習成就現況資料給教育部作為九年一貫實施成果的對照,以便將來對課程、教學及政策改進提供參考。對於導致本研究結果之因素,需藉由本研究之資料庫做進一步相關研究來發掘與推論,在保護學生個人隱私與避免不當比較的前提下,國家教育研究院籌備處明年(2009年)第二季將開放資料庫之申請,有興趣之教育相關研究人員,可連結各領域的觀點,對相關議題進行研究與探討。

關鍵詞:台灣學生學習成就評量資料庫、成就評量、九年一貫能力指標、課 程綱要

The Analysis of Taiwan Assessment of Student Achievement 2006

Chien-Ming Cheng

Assistant Researcher, National Academy for Educational Research Preparatory Office

Chin-His Chen

Chief Secretary, National Academy for Educational Research Preparatory Office

Abstract

The Taiwan Assessment of Student Achievement has been conducted for three years. The study is to present the stratified sampling results of five subjects Chinese, English, Mathematics, Social science, and Natural science administered in the five educational levels of 4th grade, 6th grade in elementary school, second grade in junior high school, second grade in senior high school, and second grade in vocational high school in 2006. The results include the descriptive statistics of the answer correct rate of competence indicators for 1st to 9th grade curricula alignment and curriculum guidelines of senior and vocational high schools as well as the statistic analysis of variables of students' background. The results can be the reference of students, teachers, and householders. In addition, the most important is that the results would be a contrast for the policy results of 1st to 9th grade curricula alignment. The results will offer to be the reference for the future of curriculum, instruction, and policy. As for the cause factors of the study results, it needs further research to investigate by the TASA database. Moreover, under the conditions of not violating students' privacy and avoiding inappropriate comparison, the Preparatory Office has provided the database for interested researchers to investigate related study.

Keywords: Taiwan Assessment of Student Achievement, TASA, database, competence indicators, 1st to 9th grade curricula alignment, curriculum guidelines; achievement assessment

壹、前言

一、建置背景與目的

依據教育基本法第九條第六項明定中央政府之教育權限為:「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,以評鑑學生的學習成效,乃中央政府之教育權限之一。

建置一完整日客觀的全國性學生學習成就資料庫,是刻不容緩的工作, 以縱貫及橫斷之研究方法來瞭解學生的學習表現,可作為日後重大教育決 策之參據。教育部於2004年10月核定本處進行台灣學生學習成就評量資料 庫(Taiwan Assessment of Student Achievement,簡稱TASA)之建置計畫, 2005年進行國小六年級國、英、數三科之施測,2006年進行國小四、六年 級、國中二年級、高中(職)二年級學生在國語文、英語文、數學、自然、 社會五科之學習成就評量(國小四年級不施測社會)。事實上,美國對於評 量該國學牛學習成就及國際性的比較接軌一向不遺餘力,且頗有成就,值得 我國借鏡效法。就國際相關研究而言,有四個知名的針對學生成就評量之整 合型研究計畫: 國際數理趨勢研究 (The Trends in International Mathematics and Science Study, TIMSS)、國家教育進展評量(National Assessment of Educational Progress, NAEP)、國際學生評量(Program for International Student Assessment, PISA)與國際閱讀發展研究(Progress of International Reading Literacy Study, PIRLS)。隨著國際互動日益密切,台灣與其他國家 (如美國、日本、英國)的接軌比較更應列入長程教育政策目標。基於上述 原因,建立台灣學生學習成就評量資料庫的目的在於(國立教育研究院籌備 處,2007):

- (一)追蹤、分析國民中小學及高中職學生在學習成就上成長變遷之趨勢, 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,藉此評估學生未來學術潛力之發

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。

二、資料庫設計內容

本資料庫以趨勢研究為主要建置骨幹,針對國小、國中、高中(職) 等三個教育階段進行施測。資料庫設計內涵採分年建置的整體規劃方式 進行。資料庫建置主要以學科屬性劃分,包含國語文、英語文、數學、自 然與社會五個科目(不含小四社會),再依其教育階段分為國小四年級、國 小六年級、國中二年級、高中(職)二年級。日後將視政策需要,不排除增 加其他年級或其他科目的可能性,以分年逐步建置方式完成。建置架構如下 圖一:



圖一 資料庫架構圖

貳、研究方法

2006年學生學習成就評量已於95年5月17~19日在全國25縣市實施,實施概況如下表1(國立教育研究院籌備處,2007)。

表1 2006年學生學習成就評量實施概況

日期	國小:5/17;國中:5/18;高中:5/19		
地區範圍	全國25縣市(包含澎湖縣、金門縣、連江縣)		
對象	利用二階段分層取樣,隨機抽取全國小四、小六、國二、高中二、高職二學生		
地點	受測學生所在學校		
科目	國語文、英語文、數學、社會、自然(學生每人施測兩科)		
作答時間	各科視考試內容及試題量不同,作答時間為40~50分鐘		

註:學生除受測二科外,另需填答受測科目之問卷及一份共同問卷。

一、評量對象

為能確保「台灣學生學習成就評量資料庫」(TASA)之建立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。

2006年施測對象為小四、小六、國二、高中二及高職二之學生,受測科目為國語文、英語文、數學、自然、社會(不含小四社會)五科,考科組合為國語文及英語文、國語文及數學、國語文及自然、國語文及社會、英語文及數學、英語文及自然、英語文及社會、數學及自然、數學及社會、自然及社會等10種。全國總施測人數為78,010人(含國語作文1,280人,英語寫說400人及自然實作200人),總校數為2,182所學校。概況如下表2:

/	,001 11	. Se H 100.1 19	c 1 xc (1)-	11/44/
施測科目	小四	小六	國二	高二
國語	8735	8678	7977	8011
英語	8498	8428	7685	7648
數學	8463	8457	7505	7631
自然	8440	8377	7629	7727
社會		8338	7652	7657
國作文	320	320	320	320
英寫說	100	100	100	100
自實作	100	100		
總計	17,428	21,499	19,484	19,599

表2 2006年各年級層施測樣本數(依科別)

二、評量工具

TASA為標準化評量,為確保試題之品質及信效度,國家教育研究院籌 備處統籌國語文、英語文、數學、自然、社會五科試題之研發,由五所教育 大學各負責一科,邀請相關學科專家、測驗專家、及各高中職、國中小學校 資深教師等組成試題研發小組,根據雙向細目表經過多次命題研習、命題、 修審題、預試、試題分析、二次修審題,最後選出優良試題,建立題庫並組 題等過程,發展出標準化成就測驗工具。

2006年學生學習成就評量各科試題內容概要如表3所示。

表3 2006年學生學習成就評量各科試題內容概要

科目,	年級	測驗內容	題目類型	試題總數	
	小四	語文應用(字詞的形、音、義及語法			
	小六	、句式的應用)、綜合性語文技巧(閱讀測驗)		36題	
國語文	國二	基本語文能力、文學文化相關知識、 閱讀測驗	選擇題		
	高中二	語文基本能力、文學知識與文章賞析、		2.4月百	
	高職二	文化知識與文化體悟、語文表達		24題	
	小四			聽20題、讀12題	
	小六	-		聽20題、讀20題	
英語文	國二	- 聽力測驗、閱讀測驗	選擇題	聽14題,讀18題	
	高中二	-		聽14題,讀22題	
	高職二	-		聽14題,讀14題	
	小四		選擇題、應用題	24題選擇題及2題應	
	小六	-		用題	
數學	國二	- 數與計算、量與實測、幾何、統計與 - 機率、代數	選擇題	18題	
	高中二	- 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			
	高職二	_			
	小六		選擇題	45題(地理12、歷 史12、公民21)	
社會	國二	- 地理、歷史、公民		45題(地理15、歷 史15、公民15)	
	高中二	-		48題(地理15、歷	
	高職二	-		史15、公民18)	
	小四	- 過程技能、科學與技術認知、科學本			
	小六	型性权能、科学典权制能加、科学本質、科學態度、思考智能、科學應用			
自然	國二	過程技能、科學與技術認知、科學態 度、思考智能、科學應用	選擇題	39題	
	高中二	高中生物、地科、物理、化學、國中 自然科			
	高職二	高職生物、物理、化學、國中自然科			

註:國語文在小學部分另有非測驗題,包括書寫與寫作試題。國中、高中與高職則為寫作 評量。英語文在小學部分另有說及寫兩部份之非選擇題。自然在小學部份,有小部份 學生進行實作測驗。

三、正式測驗題本試題分配及BIB等化設計

本研究所欲評量之目標為瞭解學生之學習成就,測驗內容含蓋廣泛、題目眾多,施測受限於學生受測時間以30~50分鐘為主的因素,並無法使每位學生都能受測到所有題目,因此本計畫中採用平衡不完全區塊(Balanced Incomplete Block, BIB)設計題本模式,其優點為試題區塊與題本的配置方式,採螺旋式排列方式,可使每一個試題區塊的施測次數相同,因此研究中將題庫中的測驗題分為幾個試題區塊(treatment),每個區塊中包含相同題數的試題,且區塊中的試題不重複。將試題區塊編製成數個題本,而每個題本中的區塊可能部分相同或完全不同,但在每個區塊出現的次數是一樣的,亦即題庫中的每個試題所受測的學生數約是相同的。其優點為試題區塊與題本的配置方式,採螺旋式排列方式,可使每一個試題區塊的施測次數相同(van der Linden, Veldkamp, & Carlson, 2004),且IRT參數估計時若採用同時估計法(concurrent calibration),則可減少等化程序所介入的額外誤差(Kim, & Cohen, 1998; Kim, & Cohen, 2002; Kolen, & Brennan, 1995; Petersen, Cook, & Stocking, 1983; Stocking, & Lord, 1983)。

四、問卷

本研究依據國民中小學九年一貫課程綱要及高中(職)課程綱要內涵,並參考各版本教科書內容及國際試題,發展標準化成就測驗工具。每一學科依其學科領域考量不同測驗內容及題型,國語科增加作文測驗,英語科增加寫說測驗。除發展各學科領域測驗評量工具,亦編製學生共同問卷及各學科問卷,以蒐集影響學生學習成就相關因素之資料。

TASA 2006年問卷共分為二大部份:第一部份為學生共同問卷;第二部份為各科問卷(包括:國語文、英語文、數學、社會及自然)。每位學生均需填寫1份共同問卷及2份各科問卷(依其施測科目而定)。

共同問卷及各科問卷以美國NELS相關研究中,對學業成就有顯著影響 變項為基礎來發展;各科問卷設計則尊重各科之意見,並授權予各科自行訂 定題目內容,而所有問卷填答方式則力求一致。共同問卷及各科問卷在經過 預試,並依預試結果調整、修改題目後,其內容大致可分為下列領域:

- ※ 學生基本資料:性別、個人於各科之喜好、交友情形,及課餘的學習情形等。
- ※ 學生家庭狀況:家庭結構、父母社經地位、父母對子女的教養態度、教育期望,及生活環境等。
- ※ 教師狀況:教師教學方式與策略、教學態度,及與學生相處情形等。

五、資料分析

本研究試題採用平衡不完全區塊(balanced incomplete block--BIB)設計,個科依試題量之不同將試題區塊編組成題本。而BIB設計最主要的考慮是能夠平衡各試題區塊的呈現次序,題本編輯以區塊為單位系統性序列安排,同時確保每一份題本施測樣本人數的適切性。

古典測驗理論一直是各種心理測驗編製與解釋之主要依據,然而其有許多理論上的缺點,因而測驗學者發展出試題反應理論,以補古典測驗理論之缺點。IRT與CTT的發展有其先後,雖立論不一,但亦有其相同之處:

(一)理論相同之處

1.試題分析

試題反應測驗理論與古典測驗理論在進行試題分析時,若試題反應測驗理論在只考慮一個參數下(亦即單參數模式),則兩種理論所推估出的難度呈現幾乎相同的極高正相關;另外在計算試題的鑑別指數時,相同道理,試題反應測驗理論在僅考慮兩個參數的雙參數模式下,其推估出的鑑別度參數a_i,應會與古典測驗理論所計算的鑑別指數(如點二系列相關係數)呈現高度正相關(Lord,1980)。

2.理論的假設

如前所述試題反應測驗理論與古典測驗理論都有其假設的前提,其中 試題的「單向度」、「知道——正確假設」之前提,對兩種理論而言,都是必要的。

3.受試者的評定

對於受試者的優劣評定,雖然試題反應測驗理論較古典測驗理論來得嚴謹(例如兩受試者答對題數相同時,試題反應測驗理論亦能分辨出兩受試者之間的能力高低),但兩種理論所得的能力值及總分,具有一致性,都能有極高的正相關。

(二)理論相異點

針對試題反應理論與古典測驗理論加以比較,其中較明顯的差異有以下四點(Hambleton & Cook, 1977;王寶墉, 1995):

1.樣本的影響

古典測驗理論以反應正確受試者之百分比作為難度指標,以試題分數與測驗總分的相關係數作為鑑別度指標,這些代表試題特質的統計量數隨受試者能力水準而變,亦即這些統計量數係團體相依(group-dependent)。而試題反應理論的試題參數,如難度參數b 及鑑別度參數a,雖然經過不同群體受試者的施測,卻仍舊保持不變,也就是說,試題參數之估計不受樣本不同的影響(sample-free)。

2.試題難度的影響

古典測驗理論的觀察分數隨測驗的難度而變,亦即觀察分數係依測驗而定(test-dependent),此現象造成不同受試者接受不同試題的結果難以直接比較。而在試題反應理論中,受試者能力估計不受試題不同的影響(item-free),且不同受試者的能力值可以直接比較。

3.測量誤差

在古典測驗理論中,所有受試者接受一測驗的結果皆具有相同的測量 誤差,無法反應出受試者間的個別差異。而在試題反應理論中,測量誤差之 估計因受試者能力水準的不同而異,因此可提供不同水準受試者能力估 計值,不同的準確度評估。

4.測驗結果的解釋

在古典測驗理論中,測驗結果的解釋通常是將受試者在測驗中答對的

試題予以計分,忽視了受試者答對試題的難度有高有低。而在試題反應理論中,則考慮試題難度的訊息,因此原始得分相同者,能力估計值不見得相同。

綜合上述,古典測驗理論雖不夠嚴謹,但理論淺顯易懂,便於在實際測驗情境(尤其是小規模測驗)實施;而試題反應測驗理論雖嚴謹,但理論艱深難懂,僅適用於本大規模測驗。所以兩派測驗理論各有所長,因此本研究將截取兩理論之所長以進行資料分析。資料回收後經過讀卡後,分析前將資料重新編排,資料分析分別採古典測驗理論(classical test theory, CTT)及試題反應理論(item response theory, IRT)進行試題特性及受試者能力分析。

而題庫題目與考生能力參數的量尺化則依據試題反應理論(Item Response Theory, IRT)(Lord, 1980)將同一考科與同年段學生作答資料進行同時的校準。主要使用軟體為BILOG-MG(Zimowski, Muraki, Mislevy, & Book, 2003),其相關公式如下:

1.古典測驗理論(郭生玉,1999)

難度係數公式:

$$P_{j} = \frac{R_{j}}{N_{i}} ,$$

其中, P_j 表第j題難度, R_j 表答對第題之人數, N_j 表第題作答總人數鑑別度係數公式:

$$r_j = \frac{(\mu_j - \mu_x)}{\sigma_x} \sqrt{\frac{P_j}{1 - P_j}} ,$$

其中, r_j 表第題鑑別度, μ_j 表答對第題之受試者之測驗總分的平均數, μ_x 表所有受試者之測驗總分的平均數, σ_x 表所有受試者之測驗總分的標準差。

2.試題反應理論(Lord, 1980)

本研究中採用三參數IRT模式,其公式如下:

$$P_{j}(\theta) = c_{j} + (1 - c_{j}) \frac{1}{1 + e^{-a_{j}(\theta - b_{j})}}$$
,

其中, $P_j(\theta)$ 表能力值為 θ 之受試者第題答對率, θ 表受試者能力值, a_j 表

試題的鑑別度, b,表試題的難度, c,表試題的猜測度。

本研究以BILOG-MG估計上述參數,以自訂標準刪除不良試題後重新估計,其不良試題之規定為至少符合下列條件之一者:

- (1) 在估計過程中無法收斂者;
- (2) r;小於0;
- (3) r_i 小於0.1且 b_i 大於3;
- (4) r_i 小於0.1且 b_i 小於-3。

量尺的建置程序,乃先將試題進行參數估計,所得考生能力估計值平均數為0,標準差為1。經由線性轉換, $\theta_{\pi}=500_{g}+250$,將量尺轉為平均數250、標準差50,分數範圍介於0到500的統一量尺。

有關影響成就之相關變項之分析,以IRT估計所得的受試者之theta值為依變項,進行單因子變異數分析(Analysis of Variance; ANOVA)。

資料之信效度:測驗工具之每一試題均依據相關資料編製說明書,詳細說明教材內容的主題、範圍、教學目標、相對應之能力指標、及其相對重要性(郭生玉,1999)。因此,具有內容效度。因採BIB設計,每一位考生因應試科目年段不同,每份考卷之題數亦不同(如表3),以數學科國二、高中、高職二最少只有18題,而高中、高職二社會科最多達48題,每道試題則都有1500位考生坐答,經分析統計後,國二之數學之信度為0.82,高職二社會則達0.91,其餘介於此兩者之間,因此,此研究之測驗工具亦具有不錯之信度。

參、2006年學生學習成就評量結果分析

以下分別說明國語文、英語文、數學、社會及自然之研究結果(未包含國語文非測驗題、英語文寫作及自然科實作):

一、國語文

(一) 小四部份

就小四整體來看,學習內容通過率以「閱讀測驗」通過率最高 72.63%,其次依序為文法與修辭72.42%、句型與應用71.25%、語 音70.36%、語意64.93%,通過率最低為「語形」,不到六成,僅有 58.86%。顯示在理解及語法方面,小四學生學習成就較佳,反而是在認字 解碼方面,語形辨識的表現較差,詳見表4。

語		單項平均答對	封率(%)	分項平均答對率(%)
	音	70.36		
字詞的音、 形、義確認	形	58.86	64.72	
// 我用在中心	義	64.93	_	68.28
語法、句式的	文法與修辭	72.42	- 71.84	
應用	句型與應用	71.25	— /1.84	
綜合性語文技巧 (閱讀測驗)		72.63		72.63

表4 小四學生國語文成就之現況描述摘要

(二)小六部份

本測驗將學生語文能力依學習內容區分為「字詞的形音義確認」、「語法句式的應用」及「閱讀測驗」三大部分,以「語法句式的應用」部分平均答對率最高為70.75%,「字詞的形音義確認」部分的平均答對率最低為64.39%,詳見表5。

	<u> </u>	• • • • •			<u> </u>	
	語文成份			對率(%)	分項平均答對率(%)	
		音	60.32			
	字詞的音、 形、義確認	形	55.67	64.39	_ 67.67	
語文	// 李玄州田町	義	70.92			
應用	語法、句式	文法與修辭	65.25	70.75		
	的應用	句型與應用	73.85	- 70.75		
Ą	综合性語文技巧((閱讀測驗)	68.72	68.72	68.72	

表5 小六學生國語文成就之現況描述摘要

(三)國二部份

國二國語文之表現成就,其學習內容之表現,我們依其學習內容取材可分三大類:現代文學、古典文學及其他。各類平均答對率如表6所示。從現代文學各項答對率觀察,「現代詩歌」表現在「現代文學」的部份最好,達65.29%。其次為「現代散文」的部份,答對率為65.07%;最差為「現代語文應用」,答對率為62.42%。分析「古典文學」的平均答對率,根據表6顯示,「古典韻文」答對率最高,達68.89%,其次為「古典語文應用」答對率達66.22%。整體觀之,雖然古典文學之「古典韻文」、「古典語文應用」答對率高,但是「古典散文」的答對率只60.46%,均低於「現代文學」的部份。

學	望內容取材	單項平均答對率(%)	平均答對率(%)
現	現代詩歌	65.29	
現代文學	現代散文	65.07	64.26
學	現代語文應用	62.42	
	古典韻文	68.89	
文 古學 典	古典散文	60.46	65.19
. , ,	古典語文應用	66.22	
	其他	44.58	44.58

表6 國二學生國語文學習內容取材平均答對率一覽表

(四)高中二

分析全國高二學生國語文能力在「語文基本能力」、「文學知識與文章賞析」、「文化知識與文化體悟」三大部份及各分項能力的表現情形如下表7。整體而言,目前高二學生均能達到三大部份的能力指標,其中以「文學知識與文章賞析能力」答對率62.27%為最高,其次是「文化知識與文化體悟能力」答對率59.44%為次高、最低是「語文基本能力」答對率是57.76%。

評量內容	各項評量內容	平均答對率(%)	平均答對率(%)
	字形字音	38.19	
語文基本能力	字詞句段義	60.11	57.76
	修辭文法	64.00	-
	古典文學知識	63.04	
文學知識與 文章賞析	現代文學知識	66.55	62.27
入平貝加	文章賞析	61.00	•
	國學常識及文化知識	61.14	
文化知識與 文化體悟	體會人生修養	54.12	59.44
△1□10回1□	判斷是非價值	61.30	-

表7 高中二學生國語文評量內容表現之現況

(五)高職二

本次測驗評量內容分別有:語文基本能力、文學知識與文章賞析、文化知識與文化體悟三大項。其中以文學知識與文章賞析(47.35%)整體表現最佳,其次是文化知識與文化體悟(46.36%)、語文基本能力(44.62%)。就表8觀之,分項評量內容成就表現是:在語文基本能力部分,以字詞句段義表現最好(48.81%),其次是修辭文法(45.21%)、字形字音(28.44%);在文學知識與文章賞析部分,現代文學知識(55.18%)表現最好,其次是古典文學知識(50.96%)、文章賞析(43.73%);在文化知識與文化體悟部分,文化傳承知識(51.90%)表現最好,其次是判斷是非價值(44.00%)、體會人生修養(42.76%)。整體而言,高職二年級學生在評量內容成就表現中,現代文學知識部分表現最佳,平均通過率達55.18%;但在字形字音的平均通過率僅有28.44%。其餘各項都在42%以上。

Ē	平量內容	各項平均答對率(%)	整體表現平均答對率(%)
	字形字音	28.44	
語文基本能力	字詞句段義	48.81	44.62
	修辭文法	45.21	_
	古典文學知識	50.96	
文學知識與文章賞析	現代文學知識	55.18	47.35
	文章賞析	43.73	_
	文化傳承知識	51.90	
文化知識與文化體悟	體會人生修養	42.76	46.36
	判斷是非價值	44.00	_

表8 高職二學生國語文評量內容成就表現之現況

二、英語文

(一) 小四部份

小四英語文之表現成就,依其語文成份分為聽與讀並按九年一貫能力 指標細分,各能力指標平均答對率如表9所示。在聽的方面以「能聽辨26個 字母」的答對率最高達92.91%,而以「能聽辨問句和直述句的語調」的答 對率最低為71.05%;在讀的方面以「能辨識印刷體大小寫字母」的答對率 最高達89.64%,而以「能看懂簡易的英文標示」的答對率最低為67.73%。

表9 小四英語文能力指標答對率分析統計摘要表

語文	成份	能力指標	答對率(%)
		1-1-1能聽辨26個字母	92.91
		1-1-2能聽辨英語的語音	77.41
		1-1-3能聽辨課堂中所習得的詞彙	83.44
接		1-1-4能聽辨問句和直述句的語調	71.05
接受性		文 性	1-1-7能聽懂常用的教室用語及日常生活用語
		3-1-1能辨識印刷體大小寫字母	89.64
	讀	3-1-2能辨識課堂中習得的詞彙	77.95
		3-1-4能看懂簡易的英文標示	67.73

(二)小六部份

小六英語文之表現成就,依其語文成份分為聽與讀並按九年一貫能力指標細分,各能力指標平均答對率如表10所示。在聽的方面與小四的結果相同以「能聽辨26個字母」的答對率最高達94.96%,而以「聽懂簡易歌謠和韻文的主要內容」的答對率最低為60.87%;在讀的方面亦以「能辨識印刷體大小寫字母」的答對率最高達92.94%,也以「能看懂簡易的英文標示」的答對率最低為73.81%。

表10 小六英語文能力指標答對率分析統計摘要表

語文原	龙份	能力指標	答對率(%)
		1-1-1能聽辨26個字母	94.96
		1-1-2能聽辨英語的語音	89.68
		1-1-3能聽辨課堂中所習得的詞彙	83.12
	聽	1-1-4能聽辨問句和直述句的語調	87.75
	铠	1-1-7能聽懂常用的教室用語及日常生活用語	85.77
		1-1-8能聽懂簡單的句子	83.80
松		1-1-9能聽懂簡易的日常生活對話	72.00
接受性		1-1-10聽懂簡易歌謠和韻文的主要內容	60.87
-La		3-1-1能辨識印刷體大小寫字母	92.94
		3-1-2能辨識課堂中習得的詞彙	84.10
			3-1-4能看懂簡易的英文標示
	讀	3-1-5能辨識故事、韻文、歌謠中的常用字詞	80.15
		3-1-6能看懂簡單的句子	83.10
		3-1-9能藉圖畫、圖示等視覺輔助,閱讀並了解簡易故事及兒童短劇 中的大致內容	76.81

(三) 國二部份

藉由全國能力指標答對機率之描述性統計摘要(表11),可得知參與 本次測驗之全體考生在聽力的學習成就表現,各能力指標的答對率平均值介 於53.92~63.17,平均值最低之能力指標為「能辨識對話或訊息的主旨或 目的」,最高為「能辨識不同句子語調所表達的情緒和態度」。而在讀的 部分,各能力指標的答對率平均值介於57.65~75.55,平均值最低之能力指標 為「能了解對話、短文、書信、故事及短劇等的重要內容與情節」,最高為 「看懂常用的英文標示和圖表」。

能力指標 題數 平均答對率(%) 1-2-1能辨識簡易詩歌的節奏與音韻 3 59.61 1-2-2能辨識不同句子語調所表達的情緒和態度 4 63.17 1-2-3能聽懂日常生活對話和簡易故事 4 61.70 1-2-4能辨識對話或訊息的主旨或目的 5 53.92 1-2-5能透過視覺上的輔助,聽懂簡易影片和短劇的大致內容 5 62.70 3-2-1能辨識英文字母的連續書寫體 4 62.88 5 3-2-3能看懂常用的英文標示和圖表 75.55 3-2-5能了解課文的主旨大意 4 59.01 3-2-6能了解對話、短文、書信、故事及短劇等的重要內容與情節 5 57.65 3-2-7能辨識故事的要素,如背景、人物、事件和結局 5 60.61 3-2-8能從上下文或圖示,猜字意或推論文意 4 64.65 3-2-9能讀不同體裁、不同主題的簡易文章 5 71.54

表11 國二英語文能力指標答對率分析統計摘要表

(四)高中二

3-2-10能了解並欣賞簡易的詩歌及短劇

藉由高中二學生能力指標答對率表(表12)可得知,參與本次測驗之 全體考生在聽力的學習成就表現平均答對率(%)介於46.80~69.68。以「能 聽懂英語日常對話」的答對率最高;而以「能聽懂教師用英語所講述的課文 內容概要,以及教師針對課文內容提出的問題」的答對率最低。

全體考生在讀的學習成就表現平均答對率(%)介於48.71~73.17。以「能看懂常用的英文標示和圖表」的答對率最高;而以「能了解短文、書信、故事、漫畫、短劇及簡易新聞報導等的內容或情節」的答對率最低。其餘各能力指標答對率如表12。

4

65.15

表12 高中二英語文能力指標答對率分析統計摘要表

項目	能力指標	題數	平均答對率(%)
	1-1能聽懂教室用語	3	65.06
	1-2能大致聽懂教師用英語所講述的課文內容概要,以及教 師針對課文內容所提出的問題	4	54.88
	1-3能大致聽懂英語日常對話	3	65.65
聽	1-4能聽懂教師用英語所講述的課文內容概要,以及教師針 對課文內容提出的問題	2	46.80
	1-5能聽懂與課文主題類似或相關之會話、故事或敘述	1	61.67
	1-6能聽懂英語日常對話	4	69.68
	1-7能聽懂英語教學廣播節目	1	60.72
	1-8能大致聽懂英語影片及國內英語新聞報導的內容	2	63.85
	1-9能大致聽懂廣播指示用語:如捷運、車站、機場廣播	1	67.44
	3-1能看懂常用的英文標示和圖表	4	73.17
	3-2能看懂短文故事並了解其大意	22	54.20
讀	3-4能利用字詞結構、上下文意、句型結構及篇章組織推測 字詞意義或句子內容	7	51.30
頙	3-5能熟悉較高層次的閱讀技巧,並有效應用於實際閱讀中	1	53.44
	3-6能了解短文、書信、故事、漫畫、短劇及簡易新聞報導 等的內容或情節	5	48.71
	3-7能了解及欣賞不同體裁、不同主題之文章	5	52.61

(五)高職二

藉由高職二學生能力指標答對率表(表13)可得知,參與本次測驗之 全體考生在聽力的學習成就表現平均答對率(%)介於33.92~48.67。以「能 聽懂教室用語」的答對率最高;而以「能大致聽懂教師用英語所講述的課文 內容概要,以及教師針對課文內容所提出的問題」的答對率最低。

全體考生在讀的學習成就表現平均答對率(%)介於34.35~64.08。以「 能看懂常用的英文標示和圖表」的答對率最高;而以「能看懂短文故事並了 解其大意」的答對率最低。其餘各能力指標答對率如表13。

表13 高職二英語文能力指標答對率分析統計摘要表

項目	能力指標	題數	平均答對(率%)
	1-1能聽懂教室用語	4	48.67
	1-2能大致聽懂教師用英語所講述的課文內容概要,以及教師針對課文內容所提出的問題	6	33.92
聽	1-3能大致聽懂英語日常對話	5	38.60
	1-4能聽懂教師用英語所講述的課文內容概要,以及教師針 對課文內容提出的問題	3	48.51
	1-6能聽懂英語日常對話	3	39.14
	3-1能看懂常用的英文標示和圖表	4	64.08
讀	3-2能看懂短文故事並了解其大意	21	34.35
	3-4能利用字詞結構、上下文意、句型結構及篇章組織推測 字詞意義或句子內容	3	45.33

三、數學

(一) 小四部份

小四測量結果,依據學生表現區分為四個等級,並根據標準設定程序,以數學量尺分數200,268,345為等級分類決斷點,將學生成就水準分為未達基礎、基礎、精熟和進階等四級。具體而言,數學量尺分數在200以下者表示該學生在此次測量未達基礎水準,表現在200至268之間的學生屬於基礎水準,268至345者表現屬精熟水準,表現在345以上者屬進階水準。表14呈現不同成就水準表現描述與人數百分比統計摘要,其中基礎以下和進階水準人數比例較少(分別為18.2%及3.2%),基礎和精熟水準人數比例較高(分別為45.5%及33.2%)。

表14 小四學生數學各成就水準表現描述與人數百分比

成就水準	表現內容	人數百分比(%)	
基礎以下 (200以下)	未達下述基礎成就水準表現內容	18.2	
基 礎 (200~268)	基礎水準的四年級學生能夠運用估算和基本事實處理簡單的整數計算,展現對分數、小數的初步理解,並且能解決簡單真實情境的數學問題。達此水準的學生能夠操作(雖然不一定精準)簡單功能的計算機、尺規、以及幾何形狀。他們在開放式問題的說明通常是有限而無法提供支持性的證據。	45.3	
精 熟 (268~345)	精熟水準的四年級學生能夠進行整數的估算、計算,並且可以辨識計算結果的合理性。在分、小數的表現上具有概念性的理解,能夠解決真實情境的數學問題。能夠適切的操作四種功能的計算機、尺規、以及幾何形體。達此水準的學生可以辨識和運用適當的訊息進行問題解決,他們在開放式問題上可以組織呈現支持性的訊息並且解釋其理由。	33.2	
進 階 (345以上)	進階水準的四年級學生能夠解決複雜及非例行的問題,他們 能夠精熟地運用計算機、尺規、以及幾何圖形。這些學生可 以形成邏輯的結論並且針對答案及解題過程提供適切而充分 的解釋。	3.2	

(二)小六部份

表15呈現小六不同成就水準表現描述與人數百分比統計摘要,其中基 礎以下和進階水準人數比例較少(分別為14.3%及9.6%),基礎和精熟水準 人數比例較高(分別為34.9%及41.1%)。

表15 小六學生數學各成就水準表現描述與人數百分比

成就水準	表現內容	人數百分比(%)
基礎以下 (193以下)	未達下述基礎成就水準表現內容	14.3
基 礎 (193~248)	學生在五個數學內容領域,能呈現初步概念性及程序性的理解,並能解決單步驟的簡單情境問題。接近精熟表現水準的基礎水準學生,可以正確辨識出解決問題所需的必要及充分資訊,但開放式問題的數學表達技巧仍有待補強。	34.9
精 熟 (248~319)	學生能統整運用數學的概念與程序解決五個內容領域的複雜情境問題,在解決問題時能靈活運用基本的運算規則,精確、有效地掌握問題中數量或空間的關係,並呈現計算以外的推理技巧。	41.1
進 階 (319以上)	學生能靈活區辨及應用各種數學規則,並且能類推、統整數學的概念及原則。他們能統整應用程序性知識和概念性理解進行複雜情境的問題解決、能進行抽象思考、具有獨特的問題解決方法、並能為解題過程提供精準且周延的書面說明。	9.6

(三)國二部份

表16呈現國二不同成就水準表現描述與人數百分比統計摘要,其中基 礎以下、基礎、和精熟水準人數比例較高(分別為33.0%、29.2%、及32.1%), 進階水準人數比例最低(5.7%)。

表16 國二學生數學各成就水準表現描述與人數百分比

成就水準	表現內容	人數百分比(%)	
基礎以下 (219以下)	未達下述基礎成就水準表現內容	33.0	
基 礎 (219~270)	基礎水準的國中二年級學生在有結構的提示(例如圖示、統計圖表)下可以正確的解題。他們可以透過適當策略、工具(包含計算機、電腦、以及幾何形狀)的選擇與運用解決數學問題。達此水準的學生在問題解決中能運用基本的代數以及非正式的幾何概念。接近精熟表現水準的基礎水準學生能辨識並運用正確解題的必要及充分資訊,但在進階的數學概念與策略上仍有待強化。	29.2	
精 熟 (270~331)	精熟水準的國中二年級學生可以理解分數、百分比、小數、以及其他數學主題(例如代數和函數)的連結關係。達此水準的學生對於基礎水準的數學運算呈現完整的理解,並且能應用程序技能於實際情境中的問題解決有充分的瞭解。 此水準的學生對於問題中的數量以及空間關係的察覺和推理均相當的熟悉。除了精熟的計算能力以外,也呈現高層次的推理技巧。在統計和機率的領域,他們能理解蒐集和組織資料的歷程並且能計算、評鑑、以及溝通結果。	32.1	
進 階 (331以上)	進階水準的國中二年級學生可以展現更複雜的數學能力,他們能運用數感(number sense)以及幾何覺知考量答案的合理性,他們可以運用更高階的抽象思考進行複雜數量、統計、幾何、以及代數的問題解決。	5.7	

(四)高中二

表17呈現不同成就水準學生量尺分數的描述統計摘要,其中基礎以下 和基礎水準人數比例較高(分別為34.8%及40.2%),精熟和進階水準人數 比例較低(分別為21.9%及3.1%)。

表17 高中二學生數學各成就水準表現描述與人數百分比

成就水準	表現內容	人數百分比(%)
基礎以下 (229以下)	未達下述基礎成就水準表現內容	34.8
基 礎 (229~285)	基礎水準的高中二年級學生的在真實情境問題中能運用代數以及幾何的推理策略來解決問題。此水準的學生能辨識以語文、代數、以及圖表表徵資訊間的關係,並且能展現幾何關係的了解以及對應至測量技巧上的知識。他們對於資料的組織與呈現、以及統計圖表的閱讀有充分的掌握,也能解決習見的簡單機率問題。	40.2
精 熟 (285~344)	精熟水準的高中二年級學生能展現對於代數、統計、幾何 以及空間推理的理解。在多項式上可以進行代數的運算, 也可理解幾何關係的証明,並且能判斷真實情境問題答案 的合理性。他們能分析詮釋統計圖表呈現的資訊,也能透 過符號化、圖形化或者表格化的函數概念元素進行辨識、 判讀與問題解決。	21.9
進 階 (344以上)	進階水準的高中二年級學生能夠理解函數概念,能夠比較 運用以數值、代數、以及圖形呈現的函數性質,會運用重 要的定理處理相關的測量問題,並且能處理複雜的空間、 曲線、球體等幾何問題。他們應能運用代數、幾何、以及 統計的知識至進階的數學領域,展現高階的數學能力。	3.1

(五)高職二

表18呈現不同成就水準學生量尺分數的描述統計摘要,其中基礎以下 和基礎水準人數比例較高(分別為43.3%及35.0%),精熟和進階水準人數 比例較低(分別為19.0%及2.8%)。

表18 高職二學生數學各成就水準表現描述與人數百分比

成就水準	表現內容	人數百分比(%)	
基礎以下 (233以下)	未達下述基礎成就水準表現內容	43.3	
基 礎 (233~289)	基礎水準的高職二年級學生在真實情境問題中應能運用估算來確認解法並且決定結果的合理性。他們能運用代數以及幾何的推理策略來解決問題。此水準的學生應可辨識語文、代數、以及圖表化資訊間的關係,並且能展現出幾何關係以及對應至測量技巧上的知識。在資料的組織與呈現、以及統計圖表的解讀上,他們能夠初步的掌握,也可解決簡單的機率問題。	35.0	
精 熟 (289~360)	精熟水準的高職二年級學生能展現對代數、統計、幾何以 及空間推理的理解。在多項式上可以進行代數的運算,也 可理解幾何關係的証明,並且能判斷和辯識真實情境問題 答案的合理性。他們能夠進行分析並且詮釋統計圖表呈現 的資訊,也能瞭解符號化、圖形化或者表格化的函數概念 元素。	19.0	
進 階 (360以上)	進階水準的高職二年級學生能夠理解函數的概念,能夠比較運用以數值、代數、以及圖形呈現的函數性質。他們能運用代數、幾何、以及統計的知識至進階的數學問題中。	2.8	

四、社會

由於社會領域範圍不僅止於單一學科,而是由三個學習內容向度:地理、歷史、公民所組成,三個學習內容向度各自包含了數個至數十個能力指標,因此在題數有限的單一題本內,欲逐題、逐條完整呈現地理、歷史、公民之能力指標,再者又必須考慮課程綱要與現行教科書的內容領域比重,以及能力向度(基本、精熟、進階)之分布,有其實際的困難。因此學科專家研議以地理、歷史、公民之主題軸替代逐條單一能力指標分析。

(一) 小六部份

依據表19所列結果顯示,在社會的三大項學習內容上,以「公民」的

答題表現72.86%為最高;「地理」的答對率為66.30%;「歷史」的答對率61.77%,為三項學習內容中較低者。

在能力向度「基本」、「精熟」、「進階」三個項目中,偏重應用、 分析、評鑑、創造的「進階」能力向度(61.06%),普遍比著重記憶的「 基本」或理解的「精熟」表現差(72.71%和66.94%);大體呈現通過率隨 能力向度難度增加而遞減的現象。

			平均答對率(%)
CX2 2121	地理		66.30
學習 · 內容 .	歷史		61.77
1111 -	公民		72.86
	基本	記憶	72.71
	精熟	了解	66.94
能力		應用	
向度	進階	分析	61.06
	近日	評鑑	01.00
		創造	

表19 小六學生社會科現況描述摘要表

1.在能力指標「地理」第三階段部份

以「1-3-6描述農村與都市在景觀和功能方面的差異表現」達成率最高,有81.51%;而以「1-3-3了解人們對地方與環境的認識與感受有所不同的原因」一項最低,只有58.48%。

2.在能力指標「歷史」第三階段部分

以「2-3-2探討台灣文化的淵源,並欣賞其內涵」最高,達到63.30%; 「2-3-3了解今昔中國、亞洲和世界的主要文化特色」其達成率最低為 58.00%。

3.在能力指標「公民」第三階段部分

以「第八主題軸-科技技術與社會」的達成率最高,達到78.97%;最

特別的是「第三主題軸-意義與價值」一項,該項達成率相當低,只達到 48.65%,是所有能力指標細目中最低者。其餘各能力指標之達成率如下表 20 °

表20 九年一貫課程社會學習領域第三學習階段能力指標達成現況

內容	能力指標	平均答對率(%)
	1-3-1了解不同生活環境差異之處,並能尊重及欣賞其間的不同特色	65.79
	1-3-2了解各地風俗民情的形成背景、傳統的節令、禮俗的意義及其在生活中的重要性	61.67
	1-3-3了解人們對地方與環境的認識與感受有所不同的原因	58.48
	1-3-4利用地圖、數據、坐標和其它資訊,來描述和解釋地表事象及 其空間組織	67.74
地理	1-3-5說明人口空間分布的差異及人口遷移的原因和結果	64.82
	1-3-6描述農村與都市在景觀和功能方面的差異	81.51
	1-3-7說明城鄉之間或區域與區域之間有交互影響和交互倚賴的關係	66.55
	1-3-8比較不同生活環境的交通運輸類型	68.42
	1-3-9分析個人特質、文化背景、社會制度以及自然環境等因素對生活空間設計和環境類型的影響	62.89
	1-3-10列舉地方或區域環境變遷所引發的環境破壞,並提出可能的解決	67.21
	2-3-1認識今昔臺灣的重要人物與事件	61.47
歷史	2-3-2探討台灣文化的淵源,並欣賞其內涵	63.30
	2-3-3了解今昔中國、亞洲和世界的主要文化特色	58.00
	三、意義與價值	48.65
	四、演化與不變	66.03
	五、自我人際與群己	78.65
公民	六、生產分配與消費	65.83
	七、權利規則與人權	78.86
	八、科技技術與社會	78.97
	九、全球關連	69.42

(二)國二部份

依據表21所列結果顯示,在社會的三大項學習內容上,以「地理」的答題表現63.29%為最高;「公民」的答對率為61.90%;「歷史」的答對率53.57%,為三項學習內容中較低者。

在能力向度「基本」、「精熟」、「進階」三個項目中,偏重理解的「精熟」能力向度答對率為56.06%,比起應用、分析、評鑑、創造的「進階」與著重記憶的「基本」表現差(56.90%和63.89%)。

			平均答對率(%)
CX2 7171	地	理	63.29
學習 内容	歷	史	53.57
110	公	:民	61.90
	基本	記憶	63.89
	精熟	了解	56.06
能力 向度	進階	應用 分析 評鑑 創造	56.90

表21 國二學生社會科現況描述摘要表

1.在能力指標「地理」第四階段部份

以「1-4-3 分析人們對地方和環境的識覺改變如何反映文化的變遷」達成率最高,有93.78%;而以「1-4-1 分析形成地方或區域特性的因素,並思考維護或改善的方法」一項最低,只有56.81%。

2.在能力指標「歷史」第四階段部分

以「2-4-1 認識臺灣歷史(如思想、文化、社會制度、經濟活動與政治 興革等)的發展過程」最高,達到57.41%;「2-4-5 比較人們因時代、 處境、角色的不同,所做的歷史解釋的多元性」,該項達成率相當低,只 達到38.19%,是所有能力指標細目中最低者。

3.在能力指標「公民」第四階段部分

以「第五主題軸一自我人際與群己」的達成率最高,達到72.69%;最 低的是「第六主題軸一生產分配與消費」一項,其達成率為50.78%。其餘 各能力指標之達成率如下表22。

表22 九年一貫課程社會學習領域第四學習階段能力指標達成現況

內容	能力指標	平均答對率(%)
	1-4-1 分析形成地方或區域特性的因素,並思考維護或改善的方法	56.81
	1-4-2 分析自然環境、人文環境及其互動如何影響人類的生活型態	67.31
	1-4-3 分析人們對地方和環境的識覺改變如何反映文化的變遷	93.78
	1-4-4 探討區域的人口問題和人口政策	62.37
地理	1-4-5 討論城鄉的發展演化,引出城鄉問題及其解決或改善的方法	73.11
	1-4-6 分析交通網與運輸系統的建立如何影響經濟發展、人口分布、 資源交流與當地居民的生活品質	59.99
	1-4-7 說出對生活空間及周緣環境的感受,願意提出改善建言或方案	67.31
	1-4-8 評估地方或區域所實施的環境保育政策與執行成果	92.54
歷史	2-4-1 認識臺灣歷史(如思想、文化、社會制度、經濟活動與政治興革等)的發展過程	57.41
	2-4-2 認識中國歷史(如思想、文化、社會制度、經濟活動與政治興革等)的發展過程,及其與臺灣關係的流變	54.38
	2-4-5 比較人們因時代、處境、角色的不同,所做的歷史解釋的多元性	38.19
	2-4-6 了解並描述歷史演變的多重因果關係	53.86
	三、意義與價值	70.91
	四、演化與不變	66.15
	五、自我人際與群己	72.69
公民	六、生產分配與消費	50.78
	七、權利規則與人權	56.52
	八、科技技術與社會	68.63
	九、全球關連	57.97

(四)高中二

依據表23所列結果顯示,在社會的三大項學習內容上,以「地理」的 答題表現63.69%為最高;「公民」的答對率為60.72%;「歷史」的答對率 58.45%,為三項學習內容中較低者。

在能力向度「基本」、「精熟」、「進階」三個項目中,偏重應用、 分析、評鑑、創造的「進階」能力向度(58.35%),普遍比著重記憶的「 基本」或理解的「精熟」表現差(59.15%和63.08%)。

			平均答對率(%)
CX2 7171	地理		63.69
學習 内容 .	歷史		58.45
1.1.0	公民		60.72
	基本	記憶	59.15
	精熟	了解	63.08
能力 向度	進階	應用 分析 評鑑 創造	58.35

表23 高中二學生社會科現況描述摘要表

1.在高中課程綱要「地理」部份

如以平均答對率的高低來代表課程標準的達成狀況,以「108-111地學 通論之人文地理」達成率最高,有67.48%;而以「3世界文化地理(一)」 一項最低,只有61.57%。

2.在高中課程綱要「歷史」部份

如以平均答對率的高低,來代表課程標準的達成狀況,以「212、 214、218-219臺灣史」最高,達到63.62%;「101-108中國上古至近世」達 成率較低,其達成率為52.78%。

3.在高中課程綱要「公民」部份

以平均答對率的高低,來代表課程標準的達成狀況,以「111-113社會 文化與社會化與社會變遷」的達成率最高,達到75.80%;「331-332國內政 治與國際關係」一項達成率相當低只達到50.19%,是所有課程標準分項細 目中最低者。其餘各能力指標之達成率如下表24。

內容		課程綱要	平均答對率(%)
	地學	101-107自然地理	61.65
	通論	108-111人文地理	67.48
地理	2本國地理	里	63.78
	3世界文化	比地理 (一)	61.57
	4世界文化	比地理 (二)	65.48
	101-108 中	中國上古至近世	52.78
歷史	109-110 \	211、213、215-217 中國近世至近代	54.38
	212、214、218-219 臺灣史		63.62
	3、4世界	文化歷史	59.80
	111-113 社	上會文化與社會化與社會變遷	75.80
八見	221-222 道德與法律規範		54.80
公民	331-332 國內政治與國際關係		50.19
	441-442 終	堅濟、環境保育與世界展望	54.33

表24 高中課程綱要社會學習領域課程標準達成現況

(五)高職二

依據表25所列結果顯示,在社會的三大項學習內容上,以「公民」的答題表現51.66%為最高;「地理」的答對率為49.22%;「歷史」的答對率38.84%,為三項學習內容中較低者。

在能力向度「基本」、「精熟」、「進階」三個項目中,偏重應用、 分析、評鑑、創造的「進階」能力向度(43.49%),普遍比著重記憶的「 基本」或理解的「精熟」表現差(49.97%和45.79%);大體呈現通過率隨 能力向度難度增加而遞減的現象。

表25 高職二學生社會科現況描述摘要表

			平均答對率(%)
CV+	地理		49.22
學習 內容	歷	史	38.84
111	公	·民	51.66
	基本	記憶	49.97
	精熟	了解	45.79
能力 向度	進階	應用 分析 評鑑	43.49
		創造	

1.在高職課程綱要「地理」部份

如以平均答對率的高低來代表課程標準的達成狀況,在五大分類中,以「201-209臺灣地理」達成率最高,有51.36%;「10011-10023、101-105地學通論」達成率相當低,只達到32.27%,是所有課程標準分項細目中最低者。

2.在高職課程綱要「歷史」部份

如以平均答對率的高低,來代表課程標準的達成狀況,在四大分類中,以「6000、7011-7082、8000世界史」達成率較高,其達成率為45.22%;而以「3031-3043、4011-4093中國上古至近世」最低,只達到36.41%。

3.在高職課程綱要「公民」部份

以平均答對率的高低,來代表課程標準的達成狀況,在四大分類中,達成率高於65%者僅有一項,「1011-1062、6011-6025、6042-6043、6051-6056、6082-6083社會學導論」的達成率最高,達到66.87%;而以「20001-20005、30003-30006、6031、6034-6035、6044政治學導論」最低,只達到39.17%。

表26 高職二學生課程綱要社會學習領域課程標準達成現況

內容	課程綱要	平均答對率(%)
	10011-10023、101-105地學通論	32.27
地	201-209臺灣地理	51.36
理	401-411中國地理	51.07
	500-519世界地理	42.27
	3031-3043、4011-4093中國上古至近世	36.41
歷	4085、4094、5000中國近世至近代	38.44
史	1091-1112、2011-2084、4044、4074臺灣史	36.62
	6000、7011-7082、8000世界史	45.22
	1011-1062、6011-6025、6042-6043、6051-6056、6082-6083社會學 導論	66.87
公民	20001-20005、30003-30006、6031、6034-6035、6044政治學導論	39.17
	4011-4104、6061-6066、6073、6081、6084經濟學導論	44.30
	20007、30001-30002、5011-5145、6032-6033、6041、6071-6072、6074-6076法律、道德與生活	50.10

万、自然

(一) 小四部份

在23個能力指標中以「權宜的運用自訂的標準或自設的工具去度量」 的平均答對率最高達85%,而以「由實驗的資料中整理出規則,提出結果」 的平均答對率最低只有23%,其餘各能力指標的平均答對率如下表27。

表27 小四自然能力指標平均答對率

能力指標	平均答對率(%)
1211 察覺事物具有可辨識的特徵和屬性	58.43
1222 權宜的運用自訂的標準或自設的工具去度量	85
1224 知道依目的(或屬性)不同,可做不同的分類	58
1241 由實驗的資料中整理出規則,提出結果	23

表27 小四自然能力指標平均答對率 (續)

—————————————————————————————————————	
1242 運用實驗結果去解釋發生的現象或推測可能發生的事	66.5
1251 能運用表格、圖表(如解讀資料及登錄資料)	48
2211 對自然現象作有目的的偵測。運用現成的工具如溫度計、放大鏡、 鏡子來幫助觀察,進行引發變因改變的探究活動,並學習安排觀測 的工作流程	60.29
2222 知道陸生(或水生)動物外型特徵、運動方式,注意到如何去改善生活環境、調節飲食,來維護牠的健康	49.29
2231 認識物質冷了外表特微之外,亦有性質的不同,例如溶解性質、磁性、導電性等。並應用這些性質來分離或結合它們。知道物質可因燃燒、氧化、發酵而改變,這些改變可能如溫度、水、空氣等都有關	57.29
2241 知道可用氣溫、風向、風速、降雨量來描述天氣。發現天氣會有變化,察覺水氣多寡在天氣變化裡扮演很重要的角色	65.67
2242 觀察月亮東昇曲的情形,以及長期持續觀察月相,發現月相盈虧, 具有週期性	80
2251 利用折射、色散,電池、電線、燈泡、小馬達,空氣或水的流動等來設計各種玩具。在想辦法改良玩具時,研討變化的原因,獲得對物質性質的了解,再藉此了解來著手改進	48.29
3201 知道可用驗證或試驗的方法來查核想法	74.5
3202 察覺只要實驗的情況相同,產生的結果會很相近	65
3203 相信現象的變化,都是由某些變因的改變所促成的	74
5211 相信細心的觀察如多一層的詢問,常會有許多的新發現	54.33
6211 能由「這是什麼?」、「怎麼會這樣?」等角度詢問,提出可探討 的問題	64.25
6221 能常自問「怎麼做?」,遇事先自行思考解決的辦法	65
6222 養成運用相關器材、設備來完成自己構想作品的習慣	61
6232 養成遇到問題時,先試著確定問題性質,再加以實地處理的習慣	81
7201 利用科學知識處理問題(如由氣溫高低來考慮穿衣)	67.43
7202 做事時,能運用科學探究的精神和方法	55.83
7203 能安全妥善的使用日常生活中的器具	65

(二)小六部份

在27個能力指標中以「相信現象的變化,都是由某些變因的改變所促成的」的平均答對率最高達98%,而以「由資料顯示的相關,推測其背後可能的因果關係」的平均答對率最低只有43%,其餘各能力指標的平均答對率如下表28。

表28 小六自然能力指標平均答對率

能力指標	平均答對率(%)
1211 察覺事物具有可辨識的特徵和屬性	93
1222 能權宜的運用自訂的標準或自設的工具去度量	96
1251 能運用表格、圖表(如解讀資料及登錄資料)	95
1312 察覺一個問題或事件 ,常可由不同的角度來觀察而看出不同的特徵	76
1332 由主變數與應變數找出相關關係	56
1341 能由一些不同來源的資料,整理出一個整體性的看法	43.75
1343 由資料顯示的相關,推測其背後可能的因果關係	43
2211 對自然現象作有目的的偵測。運用現成的工具如溫度計、放大鏡、鏡子來幫助觀察,進行引發變因改變的探究活動,並學習安排觀測的工作流程	97
2241 知道可用氣溫、風向、風速、降雨量來描述天氣。發現天氣會有變化,察覺水氣多寡在天氣變化裡扮演很重要的角色	75
2251 利用折射、色散,電池、電線、燈泡、小馬達,空氣或水的流動等來 設計各種玩具。在想辦法改良玩具時,研討變化的原因,獲得對物質 性質的了解,再藉些了解來著手改進	90
2321 察覺植物根、莖、葉、花、果、種子各具功能。照光、溫 度、 溼度、土壤影響植物的生活,不同棲息地適應下來的植物也各不相 同。發現植物繁殖的方法有許多種	54.93
2331 認識物質的性質,探討光、溫度、和空氣對物質性質變化的影響	48.14
2341 長期觀測,發現太陽升落方位(或最大高度角)在改變,在夜晚同 一時間,四季的星象也不同,但它們有年度的規律變化	57.14

表28 小六自然能力指標平均答對率(續)

2351 知道熱由高溫往低溫傳播,傳播的方式有傳導、對流、輻射。傳播時會因材料、空間形狀而不同。此一知識可應用於保溫或散熱上	65.36
2352 藉作樂器了解影響聲音高低的因素、音量大小、音色好壞等,知道樂 音和噪音之不同	65
2353 了解力的大小可由形變或運動狀態改變的程度來度量	89
2355 知道電流可產生磁場,製作電磁鐵,了解地磁、指北針。發現有些「 力」可不接觸仍能作用,如重力、磁力	72.75
3201 知道可用驗證或試驗的方法來查核想法	62
3203 相信現象的變化,都是由某些變因的改變所促成的	98
5311 能依據自己所理解的知識,做最佳抉擇	90.5
5313 相信現象的變化有其原因,要獲得什麼結果,需營造什麼變因	47.6
6211 能由「這是什麼?」、「怎麼會這樣?」等角度詢問,提出可探討的 問題	94
6222 養成運用相關器材、設備來完成自已構想作品的習慣	75
7201 利用科學知識處理問題(如由氣溫高底來考慮穿衣)	90.5
7202 做事時,能運用科學探究的精神和方法	95
7301 察覺運用實驗或科學的知識,可推測可能發生的事	73.22
7302 把學習到的科學知識和技能應用於生活中	49.4

(三)國二部份

在26個能力指標中以「把學習到的科學知識和技能應用於生活中」的 平均答對率最高達95%,而以「能依某一屬性(或規則性)去做有計畫的觀 察」的平均答對率最低只有32%,其餘各能力指標的平均答對率如下表29。

表29 國二自然能力指標平均答對率

能力指標	平均答對率(%)
1332 由主變數與應變數,找出相關關係	64
1412 能依某一屬性(或規則性)去做有計畫的觀察	32
1413 能針對變量的性質,採取合適的度量策略	50

表29 國二自然能力指標平均答對率(續)

1432 依資料推測其屬性及因果關係	48
1442 由實驗的結果,獲得研判的論點	42.67
1443 由資料的變化趨勢,看出其中蘊含的意義及形成概念	51.33
1452 由圖表、報告中解讀資料,了解資料具有的內涵性質	36.67
2351 知道熱由高溫往低溫傳播,傳播的方式有傳導、對流、輻射。傳播時會 因材料、空間形狀而不同。此一知識可應用於保溫或散熱上	79.75
2353 了解力的大小可由形變或運動狀態改變的程度來度量	85
2355 知道電流可產生磁場,製作電磁鐵,了解地磁、指北針。發現有些「 力」可不接觸仍能作用,如重力、磁力	57
2411 由探究的活動,嫻熟科學探討的方法,並經由實作過程獲得科學知識和 技能	55.86
2421 探討植物各部位的生理功能,動物各部位的生理功能,以及各部位如何協調成為一個生命有機體	58.43
2422 由植物生理、動物生理以及生殖、遺傳與基因,了解生命體的共同性及 生物的多樣性	43.29
2444 知道物質是由粒子所組成,週期表上元素性質的週期性	38.67
2445 認識物質的組成和結構,元素與化合物之間的關係,並了解化學反應與 原子的重新排列	47.5
2451 觀察溶液發生交互作用時的顏色變化	51
2452 了解常用的金屬、非金屬元素的活性大小及其化合物	59
2456 認識聲音、光的性質,探討波動現象及人對訊息的感受	44.71
5311 能依據自己所理解的知識,做最佳抉擇	92
5313 相信現象的變化有其原因,要獲得什麼結果,需營造什麼變因	77
6421 依現有的理論,運用類比、轉換等推廣方式,推測可能發生的事	44
6431 檢核論據的可信度、因果的關連性、理論間的邏輯一致性或推論過程的 嚴密性,並提出質疑	69.5
7301 察覺運用實驗或科學的知識,可推測可能發生的事	76.25
7302 把學習到的科學知識和技能應用於生活中	95
7401 察覺每日生活活動中運用到許多相關的科學概念	49
7403 運用科學方法去解決日常生活的問題	37.33

(四)高中二

在36個能力指標中以「由植物生理、動物生理以及生殖、遺傳與基因,了解生命體的共同性及生物的多樣性」的平均答對率最高達93%,而以「熱學」的平均答對率最低只有23%,其餘各能力指標的平均答對率如下表30。

表30 高中二自然能力指標平均答對率

單元名稱 / 能力指標	平均答對率(%)
群集和生態系	53
細胞和生物體	42
微生物的生命現象	48
植物的營養	30
植物的生殖生長和發育	37.25
動物的代謝和恆定性	41.67
動物的協調作用	51.4
動物的生殖和遺傳	53.4
地球科學的學科特性	30
岩漿活動	39.75
地層與地球歷史	52.5
大氣與海洋的成份結構	60
大氣與海水的運動	46.67
物理學簡介	55
熱	53.75
聲音	60
電與磁	55.5
能量與生命	80
力學	32.63
熱學	23
自然界的物質	48

表30 高中二自然能力指標平均答對率(續)

物質的形成及其變化	25
生活中的能源	49.5
物質的構造	42
物質的狀態	58
物質的變化	37.83
物質的性質	35
1432 依資料推測其屬性及因果關係	71.5
1443 由資料的變化趨勢,看出其中蘊含的意義及形成概念	86
2411 由探究的活動,嫻熟科學探討的方法,並經由實作過程獲得科學知 識和技能	82.5
2422 由植物生理、動物生理以及生殖、遺傳與基因,了解生命體的共同 性及生物的多樣性	93
2452 了解常用的金屬、非金屬元素的活性大小及其化合物	81
2456 認識聲音、光的性質,探討波動現象及人對訊息的感受	74
6431 檢核論據的可信度、因果的關連性、理論間的邏輯一致性或推論過程的嚴密性,並提出質疑	86
7401 察覺每日生活活動中運用到許多相關的科學概念	78
7403 運用科學方法去解決日常生活的問題	87

(五)高職二

在35個能力指標中以「由植物生理、動物生理以及生殖、遺傳與基 因,了解生命體的共同性及生物的多樣性」的平均答對率最高達82.33%, 而以「力矩與轉動平衡」的平均答對率最低只有21%,其餘各能力指標的平 均答對率如下表31。

表31 高中二自然能力指標平均答對率

	單元名稱 / 能力指標	平均答對率(%)
細胞和生物體		48

表31 高中二自然能力指標平均答對率 (續)

微生物的生命現象	44.4
植物的營養	39
動物的代謝和恆定性	46.43
動物的協調作用	55.8
物理學及物理量	39
直線運動	51
平面運動	32
牛頓運動定律	24.5
功與能	37.67
力矩與轉動平衡	21
靜止流體的力學性質	43
溫度與熱量	32.4
波動	40.67
靜電	38.5
電流	34
電流的磁效應	52.5
電磁感應	32.5
緒論	55.28
自然界的物質	43
物質的構造	40.67
物質的變化	38.5
氣體的性質	34
溶液的性質	35
反應速率與化學平衡	35.5
有機化合物	32
1432 依資料推測其屬性及因果關係	53.5
1443 由資料的變化趨勢,看出其中蘊含的意義及形成概念	68

表 31	高中二	白多	状 船 ナ	1 抬 煙	平均	 	(續)
1X .) I	回丁一		スミ ドド・ノン	1 4日 7元	- 1 2-1		(心目 /

2411 由探究的活動,嫻熟科學探討的方法,並經由實作過程獲得科學知識 和技能	69.5
2422 由植物生理、動物生理以及生殖、遺傳與基因,了解生命體的共同性 及生物的多樣性	82.33
2452 了解常用的金屬、非金屬元素的活性大小及其化合物	55
2456 認識聲音、光的性質,探討波動現象及人對訊息的感受	41
6431 檢核論據的可信度、因果的關連性、理論間的邏輯一致性或推論過程 的嚴密性,並提出質疑	72.5
7401 察覺每日生活活動中運用到許多相關的科學概念	67
7403 運用科學方法去解決日常生活的問題	71

肆、2006年學生學習成就表現差異分析

在性別方面,經t-test統計結果顯示,除小四自然(t=3.71)、高中二(t=3.6)、高職二(t=4.54)之數學呈顯著差異(p<.01),結果為男優於女,及小四英語(t=10.41)、小六國語(t=10.88)和英語(t=3.46)、國二(t=13.21)、高中二(t=8.55)、高職二(t=8.51)國語皆呈顯著差異(p<.01),結果為女優於男外,其餘皆無顯著差異。

在是否為原住民方面,經t-test統計結果顯示,除社會科全部(小六t=0.75、國二t=0.68、高中二t=1.08、高職二t=0.7)及自然科高中二(t=1.38)、高職二(t=0.75)不顯著外,其餘皆呈現顯著差異(p<.01),不是原住民學生顯著優於原住民學生。

在學生父親教育程度背景變項中,除社會科全部外(小六F=0.67、國二F=0.79、高中二F=0.85、高職二F=0.33),餘皆呈現顯著差異(p<.01),所有科目皆呈現父親學歷大專以上顯著優於高中職及以下,但除了小六自然及高中二國語的父親學歷為博士的小孩平均分數最高外,其餘平均分數最高的均為碩士或大專;在學生母親教育程度背景變項中,和父親教育程度背景變項的結果相似,即社會科全部不顯著外(小六F=0.46、國二F=0.84、高中二

F=0.85、高職二F=0.89),餘皆呈現顯著差異(p<.01),但在所有平均分數中最高的皆為碩士或大專。基本上父母親的教育程度是大專以上者其子女的表現皆優於高中及以下的學歷者。

在每日看電視時間方面,在小四(F=23.85)、高中二(F=4.53)、高職二(F=3.07)的自然與高中二(F=52.3)、高職二(F=22.51)的國語皆呈現顯著差異(p<.01),其結果為少於1小時者優於1~2小時者,而1~2小時者又顯著優於3小時以上者,可見這些科目與年級的學生學習成就與看電視時間呈負相關,即每天花越多時間觀賞電視的學生其學習成就越差。平均每天使用電腦時間的長短,其統計分析結果與看電視時間相同(小四自然F=96.93、高中二自然F=2.87、高中二國語F=70.37、高職二國語F=32.59,皆呈現顯著差異p<.01),但高職二自然不顯著。

而每天在家閱讀或作功課時間的變項中,小四(F=87.71)、高中二(F=31.98)、高職二(F=39.93)的國語與小四(F=63.13)的自然皆呈現顯著差異(p<.01),以每天花1~2小時閱讀或作功課的學生優於3小時以上,而每天花3小時以上閱讀或作功課的學生又優於1小時以內者,由結果可知學生花在閱讀或作功課的時間應適量就好,與傳統觀念花越多時間念書,成績就會越好的觀念不同。

伍、結論

一、評量結果與資料庫之運用

臺灣學生成就評量資料庫已邁入第三年,其結果除了提供學生、教師與家長參考外,最重要的是提供具體的學生學習成就現況資料給教育部作為九年一貫實施成果的對照,以便將來對課程、教學及政策改進提供參考。對於導致本研究結果之因素,需藉由本研究之資料庫做進一步相關研究來發掘與推論,在保護學生個人隱私與避免不當比較的前提下,國家教育研究院籌備處將於2009年第二季開放資料庫之申請,有興趣之教育相關研究人員,可連結各領域的觀點,對相關議題進行深入研究與探討。

二、各科學習成就評量規劃

學習成就評量的目的在於了解目前教育政策下,各教育階段學生學習成就的現況與趨勢,評量以各科綱要與能力指標為主,評量結果不做個人、教師、校際與縣市間之評比,亦不影響學生升學,希望學生、教師、學校、家長與縣市教育局能以平常心看待之。因此,學生不需特別準備或練習,以免增加學生課業壓力。

三、資料庫建置與分析的建議

五科之雙向細目表並不一致,根據台灣學生學習成就評量資料庫建置 的目的,是要了解九年一貫課程能力指標的達成率為何?因此,在內容向 度方面可採取九年一貫之能力指標內容為主軸,另一橫軸可用Bloom的教學 目標或NAEP的概念理解、技能過程、問題解決。並考慮以現代測驗理論一 IRT的難易度、鑑別度、猜測度等來建立題庫,在組卷的渦程能考慮將來水 平等化及垂直等化,以進行同年級不同年段及同年段不同年級間的連結與 比較。在試題方面能多方面徵題,特別是中央、地方國教輔導團的成員與 深耕種子教師的加入。另外在問卷方面,也希望教育當局能提供欲了解之議 題或現象,以便加入問卷中作為背景變項,此外,亦可參考國外類似題庫如 TIMSS、NAEP或PISA等國際知名資料庫的問卷設計內容中有意義的題目做 設計,一方面與國際接軌又可了解台灣學生學習成就在國際間的表現狀況。 針對問卷中的所有背景變項皆能做統計分析,以觀察變項變化趨勢,並且與 NELS (The Tested Achievement of the National Education Longitudinal Study of 1988 Eighth Grade Class)相同方式呈現effect size的大小。未來在國語文、英 語文能擴大寫作與口說部份,在自然科能仿造PADI的模式建立實作,數學 則希望能開發計算與證明題。

資料庫之建置乃教育之基礎工作,在歐美等國家已實施數十年,其目 的在提供資料給教育研究者研究用,藉由長期資料蒐集與研究結果來了解台 灣教育現況並促進各方面議題的討論與國際連結比較,作為教育當局政策制 定之智庫與實施成果之檢定。

參考文獻

- 王寶墉(1995)。現代測驗理論。臺北市:心理出版社。
- 郭生玉(1999)。心理與教育測驗。台北縣:菁華。
- 國立教育研究院籌備處(2007)。2006**工作成果報告書**。台北:國立教育研究院籌備處(2007)。
- Hambleton, R. K., & Cook, L. L. (1977). Latent trait models and their use in theanalysis of educational test data. *Journal of Educational Measurement*, 14, 75-96.
- Kim, S., & Cohen, A. S. (1998). A comparison of linking and concurrent calibration under item response theory. *Applied Psychological Measurement*, 22, 131-143.
- Kim, S., & Cohen, A. S. (2002). A comparison of linking and concurrent calibration under graded response theory. *Applied Psychological Measurement*, 26, 25-41.
- Kolen, M. J., & Brennan, R. L. (1995). *Test equating methods and practices*. New York: Springer Verlag.
- Lord, F. M. (1980). *Applications of item response theory to practice testing problems*. Hillsdale, N. J.: Erlbaum Publishers.
- Petersen, N. S., Cook, L. L., & Stocking, M. L. (1983). IRT versus conventional equating methods: A comparative study of scale stability. *Journal of Educational Statistics*, 8, 137-156.
- Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, 7, 201-210.
- van der Linden, W.J., Veldkamp, B.P., & Carlson, J.E. (2004). Optimizing balanced incomplete block designs for educational assessments. *Applied Psychological Measurement*, 28, 317-331.
- Zimowski, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (2003). *BILOG-MG: Multiple-group IRT analysis and test maintence for binary for binary items*. Mooresvilk IL: Scientic Software.