Bulletin of Testing and Assessment



張 元

2011年6月 第九期

研究報告

● 大學入學考試地理科試題結構之研究	李明燕
● 邏輯思考與創造能力在大學入學考試英文考科中之評量	葉錫南
● 99課綱與學科能力測驗自然考科相關議題之研究	吴國良 陳泰然
● 地球科學動畫試題的設計與應用	吳皇慶 張俊彦
● 高中生英語聽力測驗之可行性探究	游春琪

技術報告

● 歷史試題命製的基本觀念

€●大學入學考试中心(臺灣)

大學入學考試地理科試題結構之研究

李明燕

彰化師範大學

摘要

歷年大學指考試題是高中教學的重要參考,也是未來命題方向的基礎。本 文採用以結構觀點,探討指考9年地理科試題的結構,研究方法包括文本分析 法及試題反應理論,文本分析法主要運用在分析試題的結構類型,試題反應理 論分析法則用於檢視常見試題類型的難度。根據本研究對大學指考地理科進行 了九年的試題分析,研究者發現:(一)短句型試題最爲常見,(二)各類型 結構的試題難度和鑑別度與測驗內容的思考層次有關,(三)專用概念的各種 考題爲高鑑別度試題(特別是中間能力考生),一般概念兼短句型者爲簡單的 低鑑別度試題;(四)大比例地圖與衛星影像類的試題數量不多,但難度與鑑 別度適中,宜積極開發;(五)男性考生較擅長操作型概念與圖形試題,女生 較擅長文字描述型試題,故每一次的大學入學地理科考試應包括男女考生擅長 的試題。

關鍵詞:試題結構、試題難度、試題反應理論、試題特徵曲線

李明燕,彰化師範大學地理系副教授

1 🗐

A Study on the Structure of Geography Test Items of the College Entrance Examination

Ming-Yen Lee

National Changhua University of Education

Abstract

Past Geography tests of the college entrance examination has been regarded as an important reference for high school teaching, and a basis of future test design. This article adopts a structural viewpoint to analyze each test item of the Geography test in the past nine years. The research methods include textual analysis and item response theory. While textual analysis is used to analyze the structural patterns of the test items, item response theory is applied to examine the difficulties of high-frequency test items. Based on the analysis of the Geography test from 2002 to 2010, the researcher found that:

- 1. Short-sentence test items most frequently appear in the tests in the past 9 years.
- 2. Difficulty level and discrimination level of the test items are highly related with the thinking level of test content.
- 3. Test items dealing with general concepts through short sentences are simple and low discriminating, while items involving special concepts are high discriminating (especially for mid-level students).
- 4. Test items using large-scaled maps and satellite images appear scarcely, yet show a moderate degree of difficulty and discrimination. Therefore, they should be more frequently developed in the future.
- 5. Male students perform better on items of operative concepts and graphics. However, female students are more versed in descriptive items. Therefore, the annual Geography test of college entrance examination should include items that both male and female students are good at.

Keywords: tem structure, item difficulty, item response theory, item characteristics curve

Ming-Yen Lee, Associate Professor, Department of Geography, National Changhua University of Education

壹、前言

各項考試照理都要經過試題設計、試卷編製、施測與測後分析檢討四大階 段,但各階段執行的狀況深受考試定位的影響。如當前許多出版社除編輯中小 學教科書外,也出版參考書、題庫與測驗試卷,中小學教師即可省略試題設計 或試卷編製這兩個階段,此法雖可節省教師試前的作業時間,但易受制於測驗 內容是否符合教學目標。近年雖有研究以知識地圖和學習概念能量計算開發一 套教師出題的教育評量工具,讓教師能依照選定的教學範圍,建構具學習概念 權重及概念間關連強度的知識地圖,以輔助教師進行出題(張駿淵,2008)。 該研究立意雖佳,但因該系統著重在挑題,從既有的題庫中選擇試題,如果題 庫的試題品質不佳,所組成的試卷仍有待檢驗。由於坊間題庫的試題許多來自 現有的各種測驗,如歷年入學考試與各校校內考試,而各校校內考試尤其是模 擬考又常以入學考試試題的型式爲參考基準,即使是新開發的試題也是如此, 故入學考試試題的命題方向與題型對中學教育的影響性不可言喻。也因爲如 此,承辦大學入學考試試務工作的單位應從學科本質思索試題的品質。

自 2000 年代以來,測驗方式電腦化或網路化已逐漸在國內外的考試實施,如經濟合作暨發展組織(OECD)的 PISA、英國 OCR 的 e-Assessment、 美國 ETS 的 iBT-TOEFL 等;國內則有華語測驗推動工作委員會的「華語文寫 作能力測驗」(Test of TOP Writing)、多益測驗(TOEIC)的「口說與寫作 測驗」等(侯陳美,2008);也有將之運用於 E 化學習(Scalise & Gifford, 2006)。 為降低測驗過程可能的弊端,良好的硬體是基本配備,測驗的試題也需要題庫 系統支援,無論是否為適性測驗,題量都要達一定規模,降低猜答的機率,但 如何快速生產大量優質且難度又能控制的試題,則爲此法的瓶頸,或可藉由其 他領域的結構化概念設計試題。

1961年, Lewis Mumford 對古代城市的起源提出形式與典範,說明有些 村莊向城市過渡絕不僅僅是規模大小的變化,也包括方向與目的上的變化,並 體現在一種新型的組織中(宋俊嶺、倪文彥譯,2005);1970年代,Alexander 進一步對建築提出「設計型態」(Design Pattern)概念,認為構成要素之間的 關係會再三出現在不同的設計中,且有經驗的設計者,在創作時其腦海中有許 多設計型態。若以城市發展與建築設計的觀點類推到題庫開發,意即當試務單 位決定發展電腦測驗時,因作業流程迥異於紙筆測驗(Hermans & Arnhem, 2008),除要釐清發展方向與調整組織架構,也可採「設計型態」的概念,開 發數位試題的設計方法。

Draaijer & Hartog 即以設計型態(Design Pattern)的概念,擬作為高等教 育當前有用文獻與設計數位試題每日工作之間裂隙的橋樑,該計畫認為此法有 助測驗單位可以結合學習教材、學習目標與試題型態快速設計試題,其提出樣 例內容包括標題、前後關係、總結測驗的 KSA(Knowledge, Skills and Abilities)、非正式測驗的 KSA、型態核心、設計的時間、實現的時間、額外 的認知負荷、猜測的變化、計分規則等(Draaijer & Hartog, 2007)。1999-2000 年間,許擇基教授客座大考中心,為協助克服一年多試平行試卷編製的問題, 特提出「題殼」等觀念,其精神與設計型態雷同,後來雖以試題的間接資訊及 統計方法代之(許擇基,2000),但所提的建議仍可應用於題庫試題的開發。 不過,無論是設計型態或是題殼應用於數學的計算題大致可行,但其他學科或 其他類型的試題是否可行,則有待檢驗。

通常測驗單位每年為支應各項考試所需,每考科都要準備1-2份試卷,試 題至少要消耗50-100題量。雖然有研究發現,已測試題即使是調整選項位置, 難度也會隨之改變(Meyers & Miller, 2009),應可視為新試題,但由於國內 輿情對大學入學考試試題的要求相當高,測驗單位大都不會重覆使用舊試題, 故試題需年年添補與更新。試題生產的過程可分為單次設計與題庫建置兩大 類,二者各有優缺點。前者機動性高,試題易與社會發展脈動同步,但可能有 難度不穩的風險;後者雖讓試務單位可以掌握大量試題,但有些試題可能因課 綱修訂或與社會發展趨勢不一而被淘汰,資料庫需時時進行試題增刪作業,而

且試題難度因無法進行測試也可能不穩定。為增加題量與穩定試題難度,無論 是以何種方式生產與儲存試題,皆少不了學科教師參與試題設計,那麼「設計 型態」或「題殻」或許是使參與教師能快速設計質量兼備試題的可能方法。不 過,受到國內的入學考試是以學科為本位的影響,測驗內容主要著重學科的知 識體系,應試者的學習課程往往成爲命題的重要依據,試題的設計型態往往需 依學科特性分別探究。

諸學科中,地理是橫跨人文與自然的綜合性學科,高中教材除了描述性文 字,也有大量的圖表、照片、地圖、衛星影像圖,以及數位工具,其中後三者 少見於其他學科,故地理科的試題設計無法援引自其他科目的範例。若就試題 的基本組成,大致可分為外在型式的結構與學科內容,前者為外顯性,可視性 高,也較具客觀性;後者包括 KSA、認知負荷等,其中的 K (知識)與 S (技 能)較具外顯性,A (能力)與認知負荷的判斷較複雜,且多人對同一試題不 一定有相同的分類,通常可以就具修訂試題權責的試題設計者的分類為基準, 但這部分應是試務單位之內參資料,不宜委由外部人員處理。故可以先由試題 的外在型式的結構性入手,再漸次到學科內容,同時也可以先從過去已測試題 歸納出經驗法則,再修正為未來可能的方向。故本研究之目的有二:一是探討 地理科試題外在型式的結構性與難度,二是分析常用結構的試題特徵。

貳、研究觀點與方法

一、研究觀點~試題之基本結構

考試之目的無非是檢測受測者的知識與能力,知識包括字彙與概念兩大 類,二者並非截然分離,而是呈現某種連結,且字彙觸接的可能途徑也會影響 閱讀行為。如拼音文字系統的文字形體只攜帶字音的訊息,屬於間接觸接,從 字形到字義之間存在語音媒介,語音訊息的轉錄是字義產生的必要歷程;而源 於圖畫文字的漢字是用來記錄外界事物,早期主要以象形與指事的獨立文,後 來才有會意、假借、形聲與轉注,也因具備形音義特徵,漢字屬於直接觸接, 文字不需經過任何媒介,直接從字形到字義,其理解歷程很可能不同於拼音文 字。閱讀者如何記憶與理解由許多單詞按一定深層結構組成的句子,是認知心 理學家關心的課題。Haviland與 Clark提出新舊合同的看法,這種看法是任何 一個新的訊息都包含著一些舊的訊息,即溫故知新,如此新訊息的學習才能與 過去的知識發生連結,進而產生理解與記憶之目的(鄭昭明,1996)。設計試 題也是如此,應在考生既有的基礎知識提出相關連的問題或提供新的相關知 識,測驗考生對概念的理解能力是否具備遷移性或延展性,是否具備解決問題 的能力。

本研究認爲試題如同一般文章,不是任意組合的詞句即可構成試題,而是 有句法與結構,考生才得以思考。其基本組成可以分爲問題的主體(稱爲題幹) 與客體(稱爲解答項)兩大部分(圖1),前者由前置句與問題句組成,前置 句是提供文、圖、表、照片等形式的前置知識,讓考生進一步思考解決問題的 策略;後者可依作答的能動性,分爲被動式表達與主動式表達,二者所檢測的 能力完全不同。被動式表達試題是考生在試題設計者已界定的知識框架內作選 擇,當題幹缺乏前置句,如同是命令句,此等試題若是考課本內容則是測驗考 生舊有記憶的再現;當前置句與問題句語意矛盾或無關時,考生將無從思考。 主動式表達試題則賦予考生表達意見的機會,所測的能力除舊知識與新訊息聯 結程度外,尙包括知識的重組與建構能力,並依據回答內容的限制性而有不同 的思考層次。被動性表達者如選擇題、是非題、配合題;主動性表達者如填充 題、計算題、作圖題、簡答或申論題等各類題型。大體而言,解構試題的組成 方式,結構性句法基本上是假設試題如同一般文句,有一定的句法與邏輯性, 試題設計者依命題理念,以一定的句法,詮釋學科內容,而此等詮釋方式也將 直接或間接地影響學校教學。

題幹				解答項
l				
Ι		I		
前置句(句、文、圖、表等)	+	問題句	+	答案(碼、詞、句、文、圖、表等)

圖1 試題結構(李明燕, 1996; 1997)

二、試題來源與分析方法

指定科目考試(簡稱指考)試卷的版面頁數為兩張A3(8頁A4),與傳統聯考單張A3相較,試題設計顯然可以有較多彈性,再加上85年版與72年版高中課程大幅調整,本研究僅以指考歷年試題為分析樣本,即2002年至2010年九個年度的試題,其中2003年雖有為敏督莉颱風受影響的考生舉辦另一次考試,但因考生人數少,本研究未將之納入討論。由於地理科試題有選擇題與非選擇題兩大類,非選擇題由二至三大題、每大題又有2-5個子題組成,但大考中心公布或抽樣的資料皆是大題分,無法從中推算出各子題的作答情形,故本研究主要探討選擇題部分。分析的試題資料包括題型的分類與考生作答資料,題型即依據圖1的架構將每個試題拆解為題幹(總題幹與子題幹)與解答項(選項),再依其型式特徵分為文、句、詞、碼、數、圖、地圖、表、照片等類;考生作答資料除大考中心公布全體考生的難度,尙包括5,000筆考生作答的抽樣資料。

由於指考性質為高中階段的總結性評量,且為多數高中參與的考試,本研 究視各年度全體考生程度相似;由於試卷難度受到作業保密性的影響,大考中 心無法以預試取得預估值,雖然曾有研究可以共通人作為定錨(劉長瑄等, 1995),但該方法主要用於多試卷分數的等化,不是試題難度的控制;另一方 法是「試題間接資訊」,雖也可以達成多試卷分數的等化,但試題的間接資訊 仍是人為推估。顯然試題難度的穩定性端賴試題設計者的經驗及命題作業流程 各關卡的控制,故本研究在進行跨年度試題的比較,基本假設為「每年度試題 皆由有經驗者設計,試題難度主要受試題結構的影響」。

本研究之試題結構以文本分析方法,逐一剖析每個試題外在形式的基本組成,並加以歸類;之後再以統計方法,分析常用試題類型的難度與不同能力群考生的作答情形,前者以古典測驗理論計算全體考生不同題型答對率(P値);後者採試題反應理論(item response theory),以ASCAL 估算 5000 筆考生作答資料的 a (鑑別度)、b (難度)、c (猜測度),再以三參數模式計算出不同能力考生答對機率 $P(\theta)$,進而以 P與 θ 建立的試題特徵曲線(item characteristic curve,簡寫為 ICC)闡釋各試題的特色。三參數模式公式如下:(許擇基,1992)

 $P(\theta) = c_i + \frac{1}{1 + \exp[-Da_i(\theta - b_i)]}$

P(θ) 為答對率, a為試題鑑別度, 值為0~3
b為試題難度, 值為 -3~+3; θ為考生能力, 值為 -3~+3;
D值為1.7

參、研究結果與討論

指考從開始至今有九年的積累,題量達 400 題以上,已可將其視為了解指 考試題特色的資料庫。本研究依次討論歷年試題不同試題結構的題量與難度的 配置,分析常用試題結構的試題特徵曲線,進而從考生性別探討組卷應注意的 特殊試題。

一、歷年試題的組成

地理科選擇題依試題之間的關係,基本上分為單題與多題組成的題組。由 於指考為高中學生的總結性評量,試題應具備檢測考生對跨單元、跨冊統合能 力的性質,再加上地理科依據 1990 年的研究(施添福等,1990),長期取消 多選題與複選題,改以組合式的單選題代之。單題式的試題表面上雖可使試卷 檢測較多的概念,但也較不易檢驗由數個相關概念所構成的大概念,因而有題 組題。由表1的資料顯示,題組題已成為地理科主要的類型,其平均答對率不

一定較單題試題低。題組題與單題之間的差別,從外在型式來看,有總題幹者 即爲題組題,但總題幹在整組試題的角色依其與子題關係而定,有的僅是頂帽 子以特定情境將子題連結起來,敘述簡短,如9411-9412 試題;有的不只是帽 子,也提供子題解答的線索,訊息量較多,如9431-9433。若以試題設計的操 作難度來看,題組題與單題之間雖可以相互轉換,但由題組拆爲數個單題較容 易,尤其總題幹僅是帽子形式,只要各個子題題幹題意清楚,即可獨立成題; 而數個單題組合爲題組則視各題間有無共同主題,組卷者能否在短時間設計出 總題幹作爲子題間的共同題幹。故在試題設計時最好採題組題,等組卷階段再 視實際需要調整,此法除可以節省作業時間,也較容易同時編製多份平行卷。

	年度	2002	2003	2004	2005	2006	2007	2008	2009	2010	總合
	平均數(%)	68.93	68.11		45.5				56.67		65.09
單題	個數	15	19		2				9		45
	標準差	14.844	12.732		9.192				30.721		18.731
	平均數(%)	68.00	63.81	50.39	52.86	53.72	50.21	60.72	62.19	58.45	57.12
題組	個數	23	31	38	36	40	38	39	31	40	316
	標準差	23.743	20.817	21.332	18.313	20.863	23.719	22.234	25.338	20.585	22.288
	平均數(%)	68.37	65.44	50.39	52.47	53.72	50.21	60.72	60.95	58.45	58.12
總合	個數	38	50	38	38	40	38	39	40	40	361
	標準差	20.464	18.147	21.332	17.953	20.863	23.719	22.234	26.323	20.585	22.011

表1 單題與題組題選擇題的答對率與題量

試題設計時,首要考量作答方式,選擇題雖可以快速獲得考生得分,但因 無法檢測考生對特定主題的主動表達能力,故各科指考都有一定比例的非選擇 題。其比例一方面受閱卷時間與經費的影響,另一方面也因地理科論述式題型 不易閱卷,由表2的資料顯示非選擇題的比例為20%~24%,若是20%,為二 大題,每大題10分,由4-5個子題、計5個答案構成;若是24%,為三大題, 每大題8分,由3-4個子題、計4個答案構成。兩種比例看似雷同,實際上設 計操作的難度二大題較三大題難,因地理科所考的內容不外乎「what, where, why, how」等,當一大題占10%時,內容跨度需夠大,才能避免特定主題比重過高。

	年度	2002	2003	2004	2005	2006	2007	2008	2009	2010	總和
	題數	38	50	38	38	40	38	39	40	40	361
選擇	平均數(%)	68.37	65.44	50.39	52.47	53.72	50.21	60.72	60.95	58.45	58.12
	標準差	20.464	18.147	21.332	17.953	20.863	23.719	22.234	26.323	20.585	22.011
非選	題數	8	0	9	11	7	7	11	10	10	73
總和	題數	46	50	47	49	47	45	50	50	50	434

表2 選擇題與非選擇題的題數與選擇題的平均答對率

歷年選擇題的比例照理應類似,2003 年因 SRAS 之故,為避免閱卷者之 間相互傳染,當年度取消非選擇題,次年旋即恢復原比例。以試題設計的操作 難度來看,選擇題題數多,如何在有限時間量產高品質試題為主要問題;非選 擇題題數少,需儲備的量相對較少,但不易閱卷,作業難度在試務後段。從二 者之間的轉換,選擇題轉換為非選擇題較易,刪除選項再修訂提問方式即可, 但非選擇題改為選擇題則需再花時間設計選項,故資料庫建置應以選擇題為 主。

試題各組成部分的外在型式,可分符號、文字、表、圖像與綜合等五大類 十六個細類。本研究將用字量大於約100字以上者,界定為長文,小於100字 者為短句。由表3的資料發現,題幹部分以文字類試題最多,尤其是短句類; 圖像類次之,以圖與地圖較多,衛星影像與照片較少,圖像類應是當代地理科 非常重要的教學工具,但歷年試題並不多見,其因應與闡場印刷設備與成本有 關;選項部分的類別與題幹相較,僅有三大類中的六個細類,並集中在詞、代 碼、短句三類,這些類別最大特色是用字量少。若以圖1的概念討論總題幹、 子題幹與選項的關係,組合的類型更為多元(如附錄一),但題組以「短句+ 短句+短句」、「短文+短句+詞」、「短文+短句+代碼」的出現頻率最高;單 題的組合亦雷同。

總題幹、子題幹與選項之間應以何種組合爲佳,並無定論。通常試題設計

初期,設計者不一定會刻意精簡子題的用字量,但等到設計後期即需重整各子 題間的內容,相關者或共有的部分抽至總題幹供各子題使用,子題幹也得以清 楚提問,這也是總題幹的類型多樣,而子題幹何以短句居多的關鍵。若以設計 操作難度來看,總題幹最難設計,除考驗設計者的命題理念、知識廣度與撰稿 能力外,亦是檢覈資料庫對試題的管理方式,是單題存取?還是整組存取?若 是後者,一有失誤,整組即應淘汰,長文、地圖、表等類總題幹的投入量相對 較高,但也是最具延展性,可從中再產生新子題。除題量外,不同類型試題的 答對率也有所差異,但題數太少者,本文略之;其中總題幹爲長文地圖綜合類 試題的答對率最高;子題幹則以圖類試題較高,文字類的短句與長文的答對率 類似;選項因題數大都集中於符號與文字類,代碼試題通常是由多選題變異的 組合式單選題,因代碼組合本身具有一定的猜測度,答對率較高。

	*五五川		總題幹			子題幹		選項			
	親型	平均數	個 數	標準差	平均數	個 數	標準差	平均數	個 數	標準差	
尔贴	代碼							60.44	108	22.21	
11500	數字							48.31	16	22.93	
	罰							58.28	127	21.32	
文字	短句	57.45	134	22.54	58.08	312	22.052	58.11	104	21.58	
	長文	59.2	54	21.49	57.43	21	22.816				
表	表	52	15	23.58	53.25	4	25.065				
	画	54	41	22.23	64.22	9	23.684	43.6	5	30.16	
国佑	衛星影像	55.25	4	8.62	41	2	14.142				
回诼	地圖	54.57	30	21.52	56.5	8	21.428				
	照片				89	1		17	1		
	圖表	55.67	3	44.11							
	長文照片	35	3	13.08							
圖表	長文表	54.5	6	28.52	50	2	7.071				
义子 綜合	長文圖	58.89	9	21.89	86	1					
	長文地圖	70	14	19.90	54	1					
	長文地圖表	65	3	17.69							
	總和	58.12	316	22.01	58.12	361	22.011	58.12	361	22.01	

表3 試題之總題幹、子題幹與選項類型之答對率(%)

11 重

二、常用結構的試題特徵曲線

在總題幹、子題幹與選項五大類十六個細類,試題類型多達 70 種組合, 爲避免分析零碎化,本文以文字類的短句、長文以及圖像類試題進一步討論。 (一)文字類的短句試題

短句為試題總題幹與子題幹最重要的組成,對試題的影響性也最大,但其 難度是否會因其用字量少而簡單,由圖2、3、4、5四幅資料來看,單題簡短 的題幹與選項確實沒有出現連高能力群考生都不易作答的情形,其中詞類選項 的鑑別度也較短句者為佳,其因詞類多為專有名詞,若屬「專用概念」,考生 需清楚明白其意義,才易答對,如9102 試題以「建材與蛇丘地形考冰河的作 用力」,試題簡潔,又與地方居民生活素材結合,使試題難度適中又具鑑別度; 若是地名或國名,考的是「一般概念」,難度雖較低,但又可區別低能力群考 生,如9801 試題所考的「國界」。選項為短句者,若是考現象或事物的描述 式,難度較低,如9205 試題「地圖地理網格的功能」;若是考現象的因果關 係,難度雖較高,但鑑別度也高,如9201「山坡地若退耕還林,河水的變化」, 其試題特徵曲線與9102 試題精準型概念類似。



當試題爲題組題時,以短句爲主的試題,選項爲短句者多數較易,如圖4; 但若是考概念的推論,難度亦高,如9611 試題「都會區跨域合作的項目」; 或是現象的原因,如9634 試題「焚林整地的原因」。若是選項爲詞者,型式

雖與單題類似,但因多出總題幹的訊息,許多試題特徵曲線與9201 試題雷同, 高鑑度試題較多,如圖 5,這類試題若所考的概念不難,但內容不是直接取材 於課文,考生需從前置句推論,難度與鑑別度都高,如9615 試題「作物播種 季節」、9528 試題「作物產地」與9632「試題主要作物」,顯然考生對農業 的相關概念仍有待加強。



圖4 「短句+短句+短句」題型之ICC 圖5 「短句+短句+詞」題型之ICC

(二) 文字類的長文試題

一般試題設計認為「文字須力求簡明扼要,但題意要清楚明確」(郭生玉, 2004),但在教科書為一綱多本之後,題幹無疑須提供較多的前置內容,讓使 用各版本的考生據此思考,長文題也因此產生。由圖 6、7 的資料可以發現, 長文題選項以短句者較為簡單,最簡單者為 9220 試題「臺灣多地震的原因」; 當然也有難度偏高者為 9507 試題「南極發生臭氧洞的時間」。選項為代碼者, 選項往往是由多選題轉為組合式單選題,若所考的內容為常識,其難度更會因 選項的組合邏輯的可猜測性而降低,如 9318 試題「農業產銷問題」;但所考 的概念若為不同概念的連結,仍可能是高難度試題,如 9722 試題「地景與都 市郊區化過程的關係」。



(三)圖像類試題

由於地理現象繁雜,將之化繁為簡為各種圖表或解釋圖表是地理科的重要 學習內容,歷年有不少此類試題,由圖 8、9 的資料顯示,選項為詞者高鑑別 度試題較代碼者多。此類試題的圖像資料大都置於總題幹,各子題大都為短 句,若提問是考圖中要素的關係,這是平日學校教學較忽略的部分,試題難度 高,如 9637 試題「圖中甲所進行的處理」;若提問不是直接從圖解答,則試 題難度與長文題類似,如 9538 試題「油價與輕油裂解廠的區位條件的關係」。



1999 年以來,高中課程與學校教學大量使用各種圖像資料說明地理現象 與概念,並隨資訊產業的發展與政府部門對圖資管理的解嚴,人們也可以自行 使用地圖與衛星影像,指考試題也漸次引入,但或因印刷設備與成本之限,題 量仍然不多,尤其是衛星影像。若從圖 10、圖 11 的資料來看,此類試題難度

適中又具鑑別度;地圖類試題沒有特別難的試題,少數試題因圖幅受限於版 面,不易考圖內部細部的訊息,而降低試題難度,如 9125 試題「地圖表達方 式的判斷」。由於地圖與衛星影像內含豐富的地表資料,可以從中設計出許多 試題,如同長文題一樣具高度延展性,但如何將高解析度影像置於資料庫,需 電腦硬體與軟體配合,才易隨題檢索與增修,且需要有良好印刷設備,才能有 品質佳的試卷,而這也是當前地理科試題開發的最大瓶頸。



圖10 「衛星影像+短句+代碼」題型之ICC 圖11 「地圖+短句+代碼」題型之ICC

三、試題特徵曲線的性別差異

國內大學入學考試是不分性別,每位高中畢業生或同等學歷者都有資格參與,但某些能力男女先天不太一樣,試務應公平面對所有考生。在國中基測有研究發現全體受試的男生在數學科有關問題解決的表現略高於女生(盧雪梅等,2008),而女生在國文科和英語科的性別差異組型非常相似且顯著高於男生(盧雪梅,2007),指考試題是否也有如此傾向,應有專文另行討論。本文僅由歷年平均分數與試題之圖文比對試題特徵曲線。

由表 4 資料顯示,男女生平均分數的波動類似,但兩性間的差異於 2004、 2007、2010 三年達顯著水準,整體而言女生表現較佳,但 2010 年女生的分數 較男生略低。由圖 12 各試題的特徵曲線,可以發現男生與女生的差異主要在 中間能力者,且多數試題對中高能力者較具鑑別度;最為特殊的是 9932 試題 女生的曲線呈「S」型,中能力以上者的表現超越男生。若以各年試題結構類 型的分配來看(附錄二),2010 年的文字與圖像的使用量並沒有特別多,但 多幅圖像是 2005 年引入高中課程較複雜的大比例尺地形圖與衛星影像,是高中教學較不熟悉的部分。比對試題所測內容,可以發現男生較擅長操作型與圖形概念的試題,如 9904 試題「影像偏誤的原因」、9905 試題「自行車行經的 剖面」、9911 試題「衛星影像地形災害可能性判讀」、9916 試題「地形圖地 形災害判讀」、9913 試題「GIS 空間資料與屬性資料連結的計算」。女生照理 較擅處理長文型試題,如 9413 試題「消費與服務業的區位因素」;9940 試題 「水資源利用方式」,總題幹的用字量多,考生要耐心閱讀,男生的表現就不 如女生;但如果試題需運用推論能力,則男生表現較佳,如 9908 試題「影響 土壤肥沃度的原因」、9920 試題「近年台商回流的原因」、9922 試題「城市 區域發展的共同特徵」、9934 試題「埃及斑蚊分布擴大的原因」;如果圖形 試題需地方知識記憶時,高能力群女生的表現又較男生佳,如 9932 試題「歐 洲產業活動與能源消費結構的關係」。

*	領別	2002	2003	2004^{*}	2005	2006	2007^{*}	2008	2009	2010*
	平均數	52.71	57.03	40.12	36.43	40.12	38.87	48.69	49.82	45.82
男生	個數	1927	1969	1913	1981	2016	1970	1914	2016	2000
	標準差	16.443	20.426	16.005	15.671	16.765	15.288	17.239	16.414	16.476
	平均數	53.60	58.02	42.23	36.70	40.72	39.95	49.62	50.08	44.32
女生	個數	3073	3031	3087	3019	2984	3030	3086	2984	3000
	標準差	15.945	19.127	15.451	15.230	15.699	14.588	16.660	15.763	15.498

1.2

表 4 2002-2010 年指考地理科男女平均分数

說明:*代表性別差異達顯著水準







9905

考試學刊 第九期 2011.06 Bulletin of Testing and Assessment



圖12 2010年部分試題之ICC

考試學刊 第九期 2011.06 Bulletin of Testing and Assessment

肆、結論

長久以來,測驗領域對試題的研究大都著重測驗結果的統計分析,鮮少從 結構性討論試題的特色。本研究以語文的句法提出「試題結構」觀點,並以歷 年指考地理科試題爲分析對象,再佐以考生作答情形。結果發現如下:

地理科試題結構將試題分爲總題幹、子題幹與解答項,依試題的外在型式 可分為五大類十六小類組合的 70 種組合,但不是每類試題都平均出現,經常 使用的是少數類型,而且最常使用字量較少的短句型試題。此類試題雖可以有 效測驗考生的能力,設計所耗時間也較少,試卷編製也最容易,但所測內容極 易流於片斷性知識,故資料庫不宜僅有短句型試題。長文試題的內容豐富且具 延展性,又可檢測考生閱讀與理解能力,適合作爲資料的儲備試題,但因選材 費時費力,設計不易,許多設計者皆視爲畏途,宜有計畫地開發。

編製試卷時,在版面有限的情況,不宜全部為短句型試題,可以將長文、 圖、表、照片、地圖以及衛星影像等類資料設計為總題幹,子題幹再以短句提 問,且不同試題的選項類型亦需多元化,才易檢測考生的多元能力;且應考量 兩性不同的特質,有圖形試題,也應有長文試題;有推理試題,也應有描述型 試題。目前編製試卷大都側重測驗內容的分配,其實也可從試題結構著手,但 哪些類型的試題結構才能組合成良好試卷?本研究尙未觸及,建議未來可進行 各種模擬測試。

圖文並茂是地理試卷的最大特徵, 昔日視爲機密的大比例尺地形圖與衛星 影像等地景資料, 現在已是國人認識地理環境變遷的重要工具, 是「生活地理」 的最佳教材, 也是設計地理科「地景圖像」試題的好素材, 但因印刷品質與版 面的限制, 至今試題量不多, 未來資料庫可積極開發。地景圖像試題不一定僅 限於編製地理科試卷, 若大考中心鬆綁考科, 爲地景相關學系新增「地景考 科」, 這些試題即是最佳素材, 但前提是良好印刷品質。 謝辭

感謝大考中心提供歷年指考 5000 筆考生作答資料抽樣檔;陳文德先生協助整理 資料。

參考文獻

李明燕(1996)。**八十一至八十五學年度聯考地理科試題回顧**。大學入學考試中心試 題研討會資料。未出版。台北:大學入學考試中心。

李明燕(1997)。**八十六學年度聯考地理科試題分析**。未出版。台北:大學入學考試 中心。

侯陳美(2008)。「電腦測驗」簡介。選才通訊,171。台北:大學入學考試中心。 施添福、王秋原、陳文尙、陳國川(1990)。大學入學考試地理科試題的評鑑和命題

原則。收於大學入學考試中心地理科研究小組 78 年度工作計劃報告。未出版。台 北:大學入學考試中心。

張駿淵(2008)。導入知識地圖於試題評量系統。成功大學工程科學系碩士論文。

許擇基(1992)。試題作答理論簡介。台北:中國行為科學社。

許擇基、吳家怡、李明燕、吳國良(2000)。**用間接資訊來等化測驗**。未出版。台北: 大學入學考試中心。

郭生玉(2004)。教育測驗與評量。台北:精華。

劉易斯·芒福德著,宋俊嶺、倪文彥譯(2005)。**城市發展史**。北京:中國建築工業。 劉長萱、程爾觀、蔡政豐、李明燕、吳杰蓉(1995)。**學科能力測驗等化(一):共**

通題在校成績等化比較。未出版。台北:大學入學考試中心。

鄭昭明(1996)。認知心理學。台北:桂冠。

- 盧雪梅(2007)。國民中學學生基本學力測驗國文科和英語科成就性別差異和性別差 別試題功能(DIF)分析。教育研究與發展期刊,3(4),79-111。
- 盧雪梅、毛國楠(2008)。國中基本學力測驗數學科之性別差異與差別試題功能(DIF) 分析。教育實踐與研究,21(2),95-126。
- Draaijer, S. & Hartog, R. J. M. (2007). Design Patterns for Digital Item Types in Higher Education. *E-Journal of Instructional Science and Technology*, *10(1):32*.
- Hermans, P. & Arnhem, C. (2008). The Quality Management of Large-Scale Computer Based Assessments. 34th IAEA Annual Conference, Cambridge.

- Meyers, J. L., Miller, G. E., & Way, W. D., (2009). Item Position and Item Difficulty Change in an IRT-Based Common Item Equating Design. *Applied Measurement in Education*, 22: 38-60.
- Scalise, K. & Gifford, B. (2006). Computer-Based Assessment in E- Learning: A Framework for Constructing "Intermediate Constraint" Questions and Tasks for Technology Platforms. *Journal of Technology, Learning, and Assessment*, 4(6):1-32.

考試學刊 第九期 2011.06 Bulletin of Testing and Assessment

附錄一 試題型態與答對率

單題的題型與答對率

題幹	選項	個數	平均數(%)	標準差
	代碼	4	66.25	23.67
短句	詞	11	60.55	16.08
	短句	8	73.88	13.93
	代碼	4	77.00	14.86
長文	詞	3	68.00	6.08
	画	1	13.00	
長文 表	代碼	1	45.00	
長文 圖	代碼	1	86.00	
長文 地圖	代碼	1	54.00	
表	匾	3	65.33	8.15
नि	代碼	3	64.00	14.73
画	罰	4	64.25	32.73
地圖	詞	1	52.00	



題組題的題型與答對率

總題幹	子題幹	選項	題數	平均數(%)	標準差
		代碼	32	63.00	19.55
		數字	2	40.50	37.48
	短句	短句	44	54.93	23.58
		詞	38	61.66	22.68
		照片	1	17.00	
		代碼	7	46.14	7.27
运行	長文	短句	1	70.00	
起何		所回	1	18.00	
	長文 表	代碼	1	55.00	
	照片	詞	1	89.00	
	画	代碼	1	44.00	
	冻目影佈	所回	1	51.00	
	阐生影像	短句	1	31.00	
	地圖	代碼	3	47.67	32.35
		數字	2	43.00	28.28
	竡句	短句	15	65.87	15.90
	湿り	詞	21	56.95	21.31
巨大		代碼	10	52.00	24.48
天义	巨大	詞	1	84.00	
	灰义	代碼	2	80.00	14.14
	表	數字	1	17.00	
	地圖	代碼	2	73.00	2.83
毛衣 昭臣	短句	詞	2	39.50	14.85
天文 照月	長文	代碼	1	26.00	
		代碼	2	63.00	26.87
長文 表	短句	短句	2	39.50	38.89
		詞	1	37.00	
	所回	詞	1	85.00	

題組題的題型與答對率

總題幹	子題幹	選項	題數	平均數(%)	標準差
		代碼	1	69.00	
巨力 凤	后口	數字	2	36.50	19.09
長文 圖	湿何	短句	4	58.75	23.09
		詞	2	76.50	12.02
		代碼	1	95.00	
長文 地圖	短句	詞	7	72.86	17.85
		短句	6	62.50	21.06
長文 圖 表	短句	短句	3	65.00	17.692
		代碼	3	33.67	30.44
	短句	數字	1	20.00	
表	湿的	司	5	56.80	16.66
		短句	5	61.40	22.23
	地圖	短句	1	68.00	
		代碼	14	59.71	25.44
		數字	5	60.20	13.42
圖	短句	詞	15	46.00	20.10
		短句	6	49.83	20.49
		回回	1	88.00	
国主	短句	代碼	2	50.00	60.81
	湿り	短句	1	67.00	
		代碼	2	51.50	2.12
衛星影像	短句	詞	1	68.00	
		短句	1	50.00	
		代碼	9	74.56	15.58
		數字	3	65.00	20.22
	短句	詞	10	40.90	14.59
四四		短句	6	45.17	21.27
			1	48.00	
	地圖	代碼	1	43.00	



附錄二 歷年試題總題幹、子題幹與選項外在形式類型的分配

單位:題數

			符	虎	文	字	表			昌	像				ļ	圖表	文字	·綜合	1		
年	類別	數字	(代碼	詞	短句	長文	表	照片	R	圖表	衛星影像	衛星影像照片	地圖	地圖表	長文照片	長文表	長文圖	長文圖照片	長文地圖	長文地圖表	總和
	總題幹				6	6	0		0	0	0		2		0	0	0		6	3	38
2002	子題幹				30	3	1	0	1		0		0			1	1		1		38
	選項	2	13	9	14			0	0												38
	總題幹				3	9	2		8	0	0		3		3	0	0		3	0	50
2003	子題幹				42	2	1	0	4		0		1			0	0		0		50
	選項	0	9	32	9			0	0												50
	總題幹				19	10	2		5	0	0		0		0	0	2		0	0	38
2004	子題幹				36	1	1	0	0		0		0			0	0		0		38
	選項	5	12	8	12			1	0												38
	總題幹				22	6	0		3	0	0		5		0	0	0		0	0	38
2005	子題幹				36	2	0	0	0		0		0			0	0		0		38
	選項	2	15	11	10			0	0												38
	總題幹				14	5	5		8	0	2		4		0	0	0		2	0	40
2006	子題幹				35	2	0	0	1		0		1			1	0		0		40
	選項	2	11	13	12			0	2												40
	總題幹				28	0	0		5	0	0		5		0	0	0		0	0	38
2007	子題幹				34	3	0	0	0		0		1			0	0		0		38
	選項	1	10	11	16			0	0												38
	總題幹				12	3	3		4	0	0		8		0	6	3		0	0	39
2008	子題幹				37	0	0	0	1		0		1			0	0		0		39
	選項	2	13	16	7			0	1												39
	總題幹				18	3	3		2	3	0		0		0	0	2		0	0	40
2009	子題幹				31	6	1	0	2		0		0			0	0		0		40
	選項	0	12	17	10			0	1												40
	總題幹				12	12	0		6	0	2		3		0	0	2		3	0	40
2010	子題幹				31	2	0	1	0		2		4			0	0		0		40
	選項	2	13	10	14			0	1												40

Bulletin of Testing and Assessment

邏輯思考與創造能力在大學入學考試英文考科中之評量

葉錫南

臺灣師範大學

摘要

教育部於民國九十七年一月二十四日公布新版的我國普通高級中學課程 綱要,並於九十九學年度開始實施。就英文科而言,邏輯思考與創造能力的培 養列爲五大課程目標之一,是新綱要中很重要的一項變革。英文教科書開始納 入思考能力、創造力的素材與教學單元,教學現場也開始注重此一能力的培養; 然而,未來大學入學考試將如何因應、思考能力將如何評量,是全國英文教師、 學生以及家長極度關心的焦點。

本研究旨在探索未來英文科如何納入思考能力的評量,以呼應高中綱要的 精神。研究小組檢視歷年大學入學考試(學測與指考兩項考試)的英文試題, 經由內容分析與專家諮詢座談,發現選擇題型中的字彙、綜合測驗、文意選塡 與篇章結構試題,大多涵蓋課程綱要中邏輯思考與創造能力中的 1-1, 1-2, 2-1 三項能力指標,而閱讀測驗則涵蓋更廣,包括 1-1, 1-2, 1-3, 2-1, 2-2, 2-3, 2-4 等 七項。非選擇題部分,句子合併或改寫能評量到 1-1, 1-2 的能力指標,而翻譯 題並不能真正評量思考能力,作文題則可能評量到所有八項的思考能力指標。 針對這些發現,本文除了討論背後可能的原因之外,並提出數項建議,作爲未 來大學入學英文考科考試命題與組卷之參考。

關鍵詞:高中課程綱要、能力指標、思考能力評量、邏輯思考、創造力、學測、 指考



葉錫南,臺灣師範大學英語系副教授,高中英文科課程綱要召集人

Assessment of Thinking Skills and Creativity in CEEC English Tests

Hsi-nan Yeh

National Taiwan Normal University

Abstract

The new curriculum guidelines for senior high school English have been put into effect since 2010. One important feature of the new guidelines is the development of logical thinking and creativity as one of the five major curriculum goals. This new component has brought changes to teaching materials and classroom instruction in the English class. Yet, one big question, which all English teachers, students and even parents hold, to be answered is whether this new component will be included in the College Entrance Examination Center (CEEC) English tests, If yes, how will it be assessed?

This study aims to answer the above questions. Through content analysis with a focus on the thinking skills listed in the curriculum guidelines, the researchers examined the English tests of CEEC's General Scholastic Ability Tests (GSAT) and Department Required Tests (DRT) over the past decade. It was found that Vocabulary, Integrative Cloze, Vocabulary in Cloze, and Discourse Structure mostly addressed Descriptors 1-1, 1-2, and 2-1 of the curriculum guidelines. Reading Comprehension, in contrast, covered the first seven of the eight descriptors. Among the writing tests, Chinese-to-English Translation was the only test which involved very little thinking or creativity. Sentence Combining/Rewriting addressed Descriptors 1-1, 1-2, and English Composition all the eight descriptors. Possible reasons for the differences in coverage of the descriptors are discussed. In the end, some suggestions are made for the construction of GSAT and DRT.

Keywords: senior high school curriculum guidelines, ability descriptors, assessment, thinking skills, creativity, GSAT, DRT

Hsi-nan Yeh, Associate Professor, Department of English, National Taiwan Normal University, Chairperson of the Committee for Senior High School English Curriculum Guidelines

壹、前言

語言是人類溝通的工具,也是思考的媒介。許多語言學家、人類學家及心理學家都相信語言會影響人類的思考方式,有的認為語言對思考的影響是視個別語言特性而定(如1921的 E. Sapir 及 1956 的 B. L. Whorf),有的則認為此影響是跨越所有的人類語言(如 L. S. Vygotsky);有的學者認為語言中字彙影響思考最鉅,有的則認為是句法;此外,語言對思考影響力的強度是強或溫和,不同學者亦有不同的看法。儘管有這些看法上的分歧,但他們皆一致肯定語言對於人類思考方式的影響(Bloom and Keil, 2001);有的學者更肯定,在外語課程中可以透過一些策略教學來培養學生的思考能力(Lin and Mackay, 2008)。

有鑑於此,教育部民國九十七年一月二十四日公布之普通高級中學英文科 課程綱要(九十九學年度開始實施,因此簡稱為英文科「九九課綱」)特別納入 思考能力的培養,將其列為五大課程目標之一,與A、B分版並列為此次課程 變革的兩大特色(教育部,2008a)。如何在英語教學中培養思考能力,對許多 教師而言,確實是一項新穎的挑戰。因此,教育部特別以綱要補充說明(教育 部,2008b)的方式,闡釋此一理念,並提供了八個思考能力的教學示例;不過, 對於如何評量學生的思考能力,此一補充文件並未提供進一步的說明。在大學 入學考試中心的支持下,本研究規劃了一年期的計畫,探究如何在英文考試中 納入思考能力的評量,並根據研究發現提出建言,以期未來學測與指考兩項大 學入學考試的英文考科能反映高中英文教學現況,開始評量思考能力,體現課 程改革的精神¹。

¹ 感謝大學入學考試中心支持此「新課網 AB 分版測驗內容與試題探究工作計畫—英文科:AB 分版與思考能力評量」(民國 98 年 9 月至 99 年 8 月)研究計畫,本文為該計畫成果的一部分。 更感謝計畫研究伙伴尤雪瑛、陳中漢、陳秋蘭、程玉秀教授以及游春琪、林秀慧研究員殫精竭 慮的分析整合資料、深入反覆討論辯證,才有此一報告中的許多試題分析示例。

貳、研究方法

根據高中英文科課程綱要,思考能力所指的是「邏輯思考、判斷與創造力」, 包含以下三項的基本能力指標與五項的進階能力指標(教育部,2008a):

1.基本能力

1-1 能把各類訊息加以比較、歸類、排序。

1-2 能根據上下語境釐清不同訊息間的因果關係。

1-3 能分辨客觀事實與主觀意見。

2.進階能力

2-1 能分析、歸納多項訊息的共通點或結論。

2-2 能將習得的原則類推到新情境中,解決問題。

2-3 能綜合現有訊息,預測可能的發展。

2-4 能評估不同資訊,提出合理的判斷或建議。

2-5 能整合、規劃相關資訊及資源,並發揮創意。

本研究旨在探究大考中心英文科未來評量高中生上述思考能力之可行性。 欲評量思考能力,理論上有兩種可能性:(一)新闢一項新題型專門測驗此一能 力;(二)配合現有題型融入思考能力的評量。

在教學上,Fisher (2000)認為思考能力的培養,並不需要另外增加一項新 課程,而是潛藏在現有的各項課程中進行即可;而高中課程綱要補充說明(教 育部,2008b)所提供的八個示例也都是在現有課程中融入思考能力的培養。有 鑑於此,本研究團隊便優先思考上述第二種可能性,如果確實可以融入現有題 型實施思考能力的評量,一則呼應前述教學上的作法,一則將新理念融入舊架 構,僅略做必要的調整,可以降低對考生及老師的衝擊,提高其可行性。採取 的方式,是以綱要八項能力指標為架構,研究小組以團體討論(共計11次)的 方式,質性分析歷年試題的測驗內容,檢驗其思考能力的成分,加以分析彙整,

並召開兩次的專家諮詢座談,邀請大學英語教學、教育心理學學者及高中英文 老師,檢驗、修正分析結果,以回答下列研究問題:

一、以往試題是否涵蓋上述八項能力指標的全部?

二、各題型各涵蓋了哪些能力指標?

三、各題型涵蓋的重心是否有所不同?

研究小組規劃的架構是,如果以往試題完全未評量到思考能力,則必須訴 諸第一種可能性,開發新的評量方式。如果確實評量了思考能力,則需視上述 三個研究問題的發現,了解其涵蓋八項能力指標的幅度有多大、題型是否會決 定所評量的思考能力類型,方能進一步探究如何改進現有命題方式,落實思考 能力的評量。

參、研究發現

分析初期,於解讀綱要時發現,原 1-2 之能力指標「能根據上下語境釐清 不同訊息間的因果關係。」文字陳述方式使語意較爲狹隘,將「關係」僅限於 「因果關係」一種而已,排除了其他類型、常見的訊息關係(如重要性大小、 時間先後等),因此將此指標做比較寬廣的解讀,而予以調整為「能根據上下語 境釐清不同訊息之間的關係(如因果、重要性、時間順序等)」。

分析初期,另一項重要的議題是:究竟邏輯思考能力是「主動探索思考」 抑或「在被動接受多個選項之下所做的抉擇」也會運用到邏輯思考?換句話說, 選擇題由命題者事先規劃選項,而考生若依序思考而後作答,固然符合邏輯思 考過程;但若有些考生直接進入選項從中挑選較為合理者,是否就不符原先設 計評量的思考能力?經研究小組多次討論與專家諮詢座談結果決定,以正規答 題的思考流程觀之,選擇題依然提供了學生運用邏輯思考的機會。

以下依照題型,逐一簡要呈現分析結果。

一、詞彙題

分析結果發現,以往學測及指考詞彙題利用題幹情境的設計、線索的安排, 可以涵蓋思考能力 1-1, 1-2, 2-1 的能力指標。考生作答下方例 1 時必須思考, refrigerator, vacuum cleaner, dishwasher 等各項電器同屬於「家電用品」,因此選 擇 appliances。此一思考能力為「歸類」。例 2 中的 so 一字,則提供了很明顯的 線索,引導學生思考該字前後「消息來源可靠」與「不必擔心受騙」兩項訊息 間的因果關係,屬於 1-2 的能力指標。在例 3 中,考生根據第二句的訊息「短 時間內學好中文與日文兩個語言」而得到一個結論(或下一個評語)說 Nicole 是位「聰穎、厲害的」語言學習者,此種根據訊息下結論的能力,符合 2-1 的 能力指標。

例 1: We decided to buy some _____ for our new apartment, including a refrigerator, a vacuum cleaner, and a dishwasher.

(A) utensils (B) facilities (C) appliances (D) extensions (98 指考, #6)

- 1-1 能把各類訊息加以比較、歸類、排序。
- 例 2: This information came from a very _____ source, so you don't have to worry about being cheated.
 - (A) reliable(B) flexible(C) clumsy(D) brutal(98 學測, #6)

1-2 能根據上下語境釐清不同訊息間的關係(如因果、重要性、時間順序等)。

例 3: Nicole is a _____ language learner. Within a short period of time, she has developed a good command of Chinese and Japanese.

(A) convenient (B) popular (C) regular (D) brilliant

(98 學測, #4)

2-1 能分析、歸納多項訊息的共通點或結論。

二、綜合測驗

分析結果發現,以往學測及指考綜合測驗試題大多能涵蓋到思考能力 1-1, 1-2, 2-1 的能力指標。以 95 年指考第 21 題爲例,考生必須根據 complain, museum feet, the tired feeling, spending too much time 以及空格後方小孩的話"When can we sit down? What time is it?"等訊息,綜合歸納出小孩覺得參觀博物館負面的、 無聊、厭煩的感覺,此一歸納的過程屬於 2-1 能力指標。作答第 22 題時,考生 必須比較 the average visitor 與 young visitors 花在一件展覽品時間的長短,前者 時間固然短,但短暫的程度,後者甚於前者,因此必須用 even。此一比較的過 程屬於 1-1 能力指標。作答第 25 題時,因爲前一句提及 not...too many 的概念, 空格所在的本句是以一項報導作爲實例,建議一次參訪欣賞展覽品的數目,考 生必須理解到兩者關係緊密,爲主題句與示例的關係,兩者立場、角度應該一 致,因此,空格處應該是「嚴格控管數目」的 no more than,此一釐清兩個訊 息從屬關係的過程屬於 1-2 能力指標。

<u>第 21 至 25 題爲題組</u> (95指考)

Young visitors to museums often complain about having museum feet, the tired feeling one gets after spending too much time in a museum. A case of museum feet makes one feel like saying: "This is <u>21</u>. I could have done the painting myself. When can we sit down? What time is it?"

Studies of museum behavior show that the average visitor spends about four seconds looking at one object. For young visitors, the time span can be 22 shorter. Children are more interested in smells, sounds, and the "feel" of a place than looking at a work of art. If they stay in a museum too long, a feeling of boredom and monotony will build up, leading 23 to impatience and fatigue.

To <u>24</u> museum feet, try not to have children look at too many things in one visit. It is reported that young visitors get more out of a visit if they focus on <u>25</u> nine objects. One and a half hours is the ideal time to keep their eyes and minds sharp, and their feet happy!

21. (A) boring(B) difficult(C) cool(D) exciting2-1 能分析、歸納多項訊息的共通點或結論。

22. (A) almost (B) also (C) even (D) meanwhile 1-1 能把各類訊息加以比較、歸類、排序。

23. (A) efficiently (B) eventually (C) fortunately (D) permanently
1-2 能根據上下語境釐清不同訊息間的關係(如因果、重要性、時間順序等)。
24. (A) affect (B) approach (C) assure (D) avoid
2-1 能分析、歸納多項訊息的共通點或結論。

25. (A) no better than (B) no less than (C) no more than (D) no sooner than
1-2 能根據上下語境釐清不同訊息間的關係(如因果、重要性、時間順序等)。

然而,分析資料也顯示,當試題測驗目標為句法結構(如:時態、介系詞 用法等)時,雖然題幹涉及一定的訊息,但作答時不特別處理訊息間的邏輯關 係,而是著重於制式的句法觀念,邏輯思考的成分極低,因此,不列爲評量思 考能力之列。

三、文意選塡

分析結果發現,以往學測及指考文意選填試題大多能涵蓋到思考能力 1-1, 1-2,2-1 的能力指標。以 98 年指考試題為例,第 21、28 題與 2-1 能力指標有關, 第 30 題與 1-1 能力指標有關,而其餘 7 題則與 1-2 能力指標有關。就此篇選文 10 題來看,評量 1-2 能力指標占了 7 個,比例最高。類似的現象歷年一再出現, 例如 94 年學測文意選塡,其中第 35、39 題與 2-1 能力指標有關,其餘 8 題則皆 與 1-2 能力指標有關,看來此一大題最容易評量的思考能力為 1-2「能根據上下 語境釐清不同訊息間的關係(如因果、重要性、時間順序等)。」,這可能是與此

項能力指標本身著重於思考兩項訊息間的關係,關係的類別可以相當多樣(除了 1-1 的比較、歸類、排序之外,皆可納入),因此涵蓋範圍廣,出現頻率高。以 第 25 題爲例,題幹中便有很明顯的因果關係關鍵字 because 提醒考生,「這些孩 子整個孩童時期覺得自己很不重要」這個原因,會造成日後 self-esteem 較低,屬 於 1-2 的能力指標。第 28 題所在的第二段前面先談 compulsive shopping 的成因 及幼年時期的心理因素,談至本題 During adulthood...時則進展到成年時期,考 生必須根據前述多項訊息(包括成因及心理)綜合推斷,與幼年時的玩具一樣, 此時的「購買」功用在於替代「愛意」(affection),所思考的訊息關係不僅有 1-1 的類比,還涉及更多訊息的整合判斷,因此宜歸為 2-1 的能力指標。第 30 題前 文所談皆為一般飽受 compulsive shopping 之苦者,而在此句則以其中的美國人為 例,說明其占了總人口的 8.9%,此種類別(前者)與例子(後者)間的關係, 屬於 1-1 指標的「歸類」。

第 21 至 30 題爲題組 (98 指考)

Oniomania is the technical term for the compulsive desire to shop, more commonly referred to as compulsive shopping or shopping addiction. Victims often experience feelings of 21 when they are in the process of purchasing, which seems to give their life meaning while letting them forget about their sorrows. Once 22 the environment where the purchasing occurred, the feeling of a personal reward would be gone. To compensate, the addicted person would go shopping again. Eventually a feeling of suppression will overcome the person. For example, cases have shown that the bought goods will be hidden or destroyed, because the person concerned feels 23 of their addiction and tries to conceal it. He or she is either regretful or depressed. In order to cope with the feelings, the addicted person is prompted to 24 another purchase.

Compulsive shopping often begins at an early age. Children who experienced parental neglect often grew up with low <u>25</u> because throughout much of their childhood they felt that they were not important as a person. As a result, they used

toys to <u>26</u> their feelings of loneliness. Because of the ongoing sentiment of deprivation they endured as children, adults that have depended on materials for emotional <u>27</u> when they were much younger are more likely to become addicted to shopping. During adulthood, the purchase, instead of the toy, is substituted for <u>28</u>. The victims are unable to deal with their everyday problems, especially those that alter their self-esteem. Important issues in their lives are repressed by <u>29</u> something. According to studies, as many as 8.9 percent of the American population <u>30</u> as compulsive buyers. Research has also found that men and women suffer from this problem at about the same rate.

(A) support	(B) qualify	(C) affection	(D) ashamed
(E) make up for	(F) leaving	(G) turn to	(H) buying
(I) self-esteem	(J) contentment		

參考答案: JFDGI, EACHB

第30題:1-1 能把各類訊息加以比較、歸類、排序。

第 22、23、24、25、26、27、29 題:1-2 能根據上下語境釐清不同訊息間的關係(如因果、重要性、時間順序等)。

第21、28題:2-1能分析、歸納多項訊息的共通點或結論。

四、篇章結構

分析結果顯示,篇章結構試題可以評量到思考能力 1-1, 1-2, 2-1 三項能力指標。以 94 年指考試題爲例,第 45 題評量 1-1,第 41-43 題評量 1-2,第 44 題評量 2-1 能力指標。第 41 題作答時,考生必須瞭解前一句談及荷蘭的實驗,而後一句點出實驗名稱爲 Raven,本題介乎其間,句子未經轉折,必然還是在談該實驗的細節,因此會選擇選項(C);此種比較本題(尤其是 An intelligence test)與前後訊息(One of the studies 以及 The test)之間指涉(reference)關係的思考,屬於 1-2 能力指標。作答第 42 題時,必須了解到前一句 The researchers found a strong relationship between the birth order of the test takers and their scores on the

Raven test.帶出一項主題(關鍵字為 relationship between the birth order... and... scores),而本題內容則為該主題之確切細節(關鍵字為與前句重複的 Scores... birth order.),用以解釋此一研究之發現,是主題與細節之間的關係,屬於 1-2 能力指標。作答第 43 題時,考生必須根據上下語境"In 1975, Zajonc and Markus developed the confluence theory..."、"Rutherford and Sewell in 1991 tested the theory..."與本題三句之間的時間順序(包括 In 1975, since then, in 1991),以及 三句語意一致連貫關係(三句皆提到 the theory),進而推斷本題正答句子"Since then, the theory has been elaborated and even extended..."應該是居中連接上述二句,此思考過程亦屬於 1-2 能力指標。作答第 44 題時,考生必須了解空格後之 文意 On one side...及 On the other side...各陳述正、反兩面的看法,因此歸納出 空格應該為談及兩面看法的主題句 There are definitely two sides to this issue.,此 歸納的思考過程屬於 2-1 能力指標。第 45 題評量重點在於空格正答"On the other side are Rutherford and Sewell,..."與"On one side there are Zajonc and Markus,..." 之間的對比關係;此種比較兩者的思考過程,屬於 1-1 能力指標。

<u>第 41 至 45 題爲題組</u> (94 指考)

Many researchers have been interested in whether or not an individual's birth order has an effect on intelligence. One of the first studies was carried out in the Netherlands during the early 1970s. <u>41</u> The test was called the "Raven," which is similar to the I.Q. test. The researchers found a strong relationship between the birth order of the test takers and their scores on the Raven test. <u>42</u>

In 1975, Zajonc and Markus developed the confluence theory to explain the negative effect of birth order on intelligence involving the data from the Dutch. <u>43</u> However, Rutherford and Sewell in 1991 tested the theory and found no support for it. They concluded that birth order effects did not exist.

<u>44</u> On one side there are Zajonc and Markus, who state that birth order effects may be explained solely by family size and the spacing of births. With short birth intervals, increasing order of birth will be associated with lower intelligence
levels. But with long birth intervals, this pattern may be reversed. <u>45</u> They show that the confluence model does not explain any relationship between birth order and intelligence that may exist in the American data. Up to date, there is no agreement between these opposing views. And such a debate may continue for years to come.

- (A) There are definitely two sides to this issue.
- (B) Scores decreased as the family size increased and also with birth order.
- (C) An intelligence test was administered to over 350,000 Dutch males when they turned 19 years of age.
- (D) On the other side are Rutherford and Sewell, who studied more than 10,000 American high school graduates.
- (E) Since then, the theory has been elaborated and even extended to explain the positive effect of birth order on intelligence.

參考答案:CBEAD

第 41、42、43 題:1-2 能根據上下語境釐清不同訊息間的關係(如因果、重要 性、時間順序等)。

第44題:2-1能分析、歸納多項訊息的共通點或結論。

第45題:1-1 能把各類訊息加以比較、歸類、排序。

五、閱讀測驗

分析結果顯示,閱讀測驗試題可以評量到思考能力 1-1, 1-2, 1-3, 2-1, 2-2, 2-3, 2-4 七項能力指標。以 94 年指考試題為例,第 50 題很顯然的是評量考生綜 合全篇文意,從各項訊息找出共通點、甚至提高一個層次找出主旨的能力,本 題正答(D)即屬此類主旨,而其他三個選項僅為個別單一細節。凡是"What is the main message of the passage?"、"What is the best tile for the passage?"或 "What is this passage mainly about?"這類問主旨、主題的問題,皆屬於 2-1 之能 力指標。第 51 題的作答線索在第三段,考生由文章中對聽眾反應的描述 (murmuring and shuffling their feet 等),應可判斷聽眾覺得無聊(They were bored.) 而非印象深刻(They were impressed.), 加上第四段"...not exactly a factor in these people's everyday experience." 點出與聽眾生活關連不大,從這些訊息應 該可以推斷觀眾覺得無聊,是因為演講主題非其日常生活所關心者(…not their daily concern),因此選擇(A),此一推斷的思考過程符合 2-4「能評估不同資 訊,提出合理的判斷或建議」之能力指標。作答第52題時,考生應可從文章整 理歸納出,第二段的 immune system disorder.及 Creutzfeldt-Jakob disease 與第三 段的 genetic research 為作者所談及,而第四段出現的 flu 與前三個詞不同,係 出自商會主席之口,而非作者演講所提。此種整理歸類的過程,即是 1-1 之能 力指標。第 53 題評量學生對比喻用語"had them on the edge of their seats"的 理解,此用語跳脫字面上的意義,考生從第二、三段了解在演講主題過於專業、 冷僻,以致於聽眾坐立不安,而在第四段則由商會主席之口點出相反的情況, 演講主題若與聽眾生活連結,聽眾便會身體前傾坐在椅子前緣,全神貫注聆聽 演講。考生必須分析歸納此三段內容之訊息,才能找出答案,此分析歸納的過 程,即是 2-1 之能力指標。原考題並無第 54 題,因為遍尋歷年之學測、指考試 題,皆未發現有任何試題觸及 1-3「能分辨客觀事實與主觀意見。」之能力指 標,此處爲舉例說明,特別設計此題。此題四個選項皆爲與選文相關之訊息, 但其中有些屬於明確可見的客觀事實,例如聽眾在竊竊私語、在扭動足部、及 看報紙等現象,而文中提及聽眾很不自在、不安份 (restless), 是作者的感覺, 是一種主觀的看法及觀感,藉由這四個選項,讓考生區分客觀事實與主觀意見, 評量的是 1-3 之能力指標。

第 50 至 54 題爲題組 (94 指考)

Dr. Thompson was pleased. Just three months after moving to the small Midwestern town, he had been invited to address an evening meeting of the Chamber of Commerce. Here was the perfect opportunity to show his knowledge of modern medicine and to get his practice off to a flourishing start. With this in mind, the doctor prepared carefully.

On the night of his speech, Dr. Thompson was delighted to see that the meeting hall was full. After being introduced, he strode confidently to the lectern and announced his topic: "Recent Advances in Medicine." He began with a detailed discussion of Creutzfeldt-Jakob disease, a rare brain disorder that had recently been covered in the New England Journal of Medicine. Next he outlined the progress that had been made in studying immune system disorders.

Just about this time, halfway through his speech, Dr. Thompson began to sense a certain restlessness in his audience. People were murmuring and shuffling their feet. Some seemed to be glancing at a newspaper. Nevertheless, Dr. Thompson plowed on. He had saved the best for last. He quoted extensively from an article in the Lancet about genetic research, feeling sure his audience would be impressed by his familiarity with this prestigious British medical journal.

Then the speech was over. Dr. Thompson had expected to be surrounded by enthusiastic people, congratulating him and asking questions. Instead he found himself standing alone. Finally the president of the Chamber of Commerce came up to him. "Something tells me," said Dr. Thompson, "that my speech was not very successful. I can't understand it. I worked so hard to make it interesting." "Oh, it was a fine speech," replied the president. "But maybe it would have gone over better with a different audience. Creutzfeldt-Jakob disease is not exactly a factor in these people's everyday experience. You know, here we are in January. If you'd talked about ways to avoid getting the flu, you'd have had them on the edge of their seats!"

- 50. What is the main message of the passage?
 - (A) A good speaker has to be fully prepared regardless of the audience.
 - (B) A good speaker should display his learning to the audience in an enthusiastic way.
 - (C) The more a speaker wants to please the audience, the more likely he will succeed.
 - (D) The key to a successful speech is to make it meaningful and relevant to the audience.
 - 2-1 能分析、歸納多項訊息的共通點或結論。

- 51. What was the reaction of the audience to Dr. Thompson's speech?
 - (A) They were bored because the medical topics were not their daily concern.
 - (B) They did not understand him so they could only discuss among themselves.
 - (C) They were impressed by his familiarity with advanced research in medicine.
 - (D) They congratulated him on the success of the speech and asked him questions.
 - 2-4 能評估不同資訊,提出合理的判斷或建議。
- 52. Which topic was NOT mentioned in the doctor's speech?
 - (A) Genetic research. (B) Flu.
 - (C) Immune system disorder. (D) Creutzfeldt-Jakob disease.
 - 1-1 能把各類訊息加以比較、歸類、排序。
- 53. What does "had them on the edge of their seats" mean in the last sentence of the passage?
 - (A) Had them stand up. (B) Took them by surprise.
 - (C) Caught their full attention. (D) Aroused their suspicion.
 - 2-1 能分析、歸納多項訊息的共通點或結論。
- 54. Which of the following is Dr. Thompson's personal opinion rather than a fact about the audience?
 - (A) Some people were murmuring.
 - (B) Some people were shuffling their feet.
 - (C) Some people were reading newspapers.
 - (D) Some people were feeling restless.
 - 1-3 能分辨客觀事實與主觀意見。

(第54題為新編試題)參考答案:DABCD

六、句子合併/改寫

過去學測與指考雖然未考過句子合併或改寫,但句子合併與句子改寫仍是 學測可能的基本寫作能力測驗方式,檢視、分析大考中心過去寫作研究案的這 類試題,發現多數試題皆在測驗 1-2 或 1-1 的能力指標。第 1 題中,考生必須 釐清 speak English well 與 go to college in the USA 兩者間的關係,前者爲後者的 條件,當此條件缺乏時(doesn't speak English well),仍然決定進行後者,此一 釐清彼此關係的過程,合乎「1-2 能根據上下語境釐清不同訊息之間的關係(如 因果、重要性、時間順序等)」的能力指標。第 5 題中,考生必須比較後才得知 兩句中的書本數目一樣多²,而必須使用...as many books as...,此一比較過程, 符合「1-1 能把各類訊息加以比較、歸類、排序」的能力指標。

<u> 句子合併/改寫 (英文寫作能力測驗規劃研究 VI)</u>

1. (even if)

Lucy doesn't speak English well. She has decided to go to college in the USA.

1-2 能根據上下語境釐清不同訊息之間的關係(如因果、重要性、時間順序等)

5. (as...as)

My daughter read 20 books last month. Her classmate Judy read the same number of books last month.

1-1 能把各類訊息加以比較、歸類、排序

² 此題若要使思考能力的特質更為明顯,宜將第二句改為 Her classmate Judy read 20 books last month.讓學生真正的比較書本數,而不是可以借重第二句的 Her classmate Judy read the same number of books last month.得知結果。

七、翻譯

學測與指考的翻譯題是中文翻譯為英文句子,不管是單句翻譯、克漏式翻譯、或連貫式翻譯,由於題目提供的中文句子已經將各語意單位間的關係界定 清楚,考生僅需要用詞彙、文法、句構知識與能力即可作答,無須再做語意間 的邏輯推敲、思考。即使有,也是非常有限,因此,我們將其界定為未實質評 量思考能力。

八、英文作文

分析結果顯示,英文作文試題可以評量到思考能力1-1,1-2,1-3,2-1,2-2, 2-3, 2-4, 2-5 全部八項的能力指標。以 96 年學測作文四格漫畫試題為例, 在決 定四張圖片順序時,學生便可能開始運用多種思考能力,例如,將四圖依照室 内室外區隔開來,此種思考符合「1-1 能把各類訊息加以比較、歸類、排序」 的能力指標。同時,考生還會根據圖片之間可能的因果或時間先後關係,決定 圖片合理的順序,此思考過程符合「1-2 能根據上下語境釐清不同訊息之間的 關係 (如因果、重要性、時間順序等)」的能力指標。除了依照 1, 2, 3, 4 的順序 铺陳故事外,考生也能根據自己的思考邏輯,決定自己的圖片順序,編寫一個 合乎邏輯的故事。例如按第1、4、3圖的順序,女孩擅自將貓帶回家(圖1), 但小貓將客廳弄得一團亂 (圖4);其結果是媽媽出面處理,與女孩懇談達成共 識(圖3),考生應該能根據前三張圖而得到「女孩將小貓送走」(圖2)的結論。 此思考邏輯符合「2-1 能分析、歸納多項訊息的共通點或結論」的能力指標。 在此邏輯模式之下,考生選定圖 4 之後,便會針對圖 4,合理推斷、預期母親 在看到客廳慘狀之後必然會找女孩詳談;而在圖 4 之後安排圖 3 繼續故事的發 展,此一思考過程符合「2-3 能綜合現有訊息,預測可能的發展」的能力指標。 部分考生有可能進一步根據女孩的感受、媽媽的困擾、家裡的空間等相關因素, 綜合評估,提出合理的方案,建議這個家庭該如何處理這些貓,此種思考過程 符合「2-4 能評估不同資訊,提出合理的判斷或建議」的能力指標。在思考如 何串連各張圖片時,甚至在鋪陳各單張圖片的情節時,學生皆有極大的空間發 揮不同的創意,符合「2-5 能整合、規劃相關資訊及資源,並發揮創意」的能 力指標。考生在上述整個作文寫作過程中,至少涵蓋到1-1,1-2,2-1,2-3,2-4,2-5 等六項能力指標。考生在描述母女兩人談話時,不管是母親說明不該留下小貓 的理由,或女兒陳述個人喜歡小貓的感受而極力想要留下小貓,若能適時舉出 好的客觀事實加以佐證,例如醫學報導(如:貓與傳染疾病關係)、心理學研究 (如:養動物與人際關係的實驗)、生活周遭實際案例等,則可以強化個人論點, 如此善用客觀資料,則是1-3 能力指標的表現。

英文作文(96學測)

提示:請以下面編號1至4的四張圖畫內容為藍本,依序寫一篇100至120個 單詞左右的文章,文章,描述女孩與貓之間的故事。你也可以發揮想像 力,自己選定一個順序,編寫故事。請注意,故事內容務必涵蓋四張圖 意,力求情節完整、前後發展合理。



肆、討論與建議

綜合上述分析資料顯示,在選擇題部分,歷年學測與指考試題中的題型, 除了文法題,多數題目皆會相當程度地涵蓋到某個面向的思考能力,詞彙題、 綜合測驗題、文意選塡題、篇章結構題,皆可以評量到考生在高中課程綱要中 1-1,1-2,2-1方面的思考能力,而閱讀測驗試題則可以評量到1-1,1-2,1-3,2-1, 2-2,2-3,2-4 七項思考能力指標,與前述其他大題僅能評量到三項能力指標相當 不同。思考能力指標與選擇式題型的關係,大致可以簡化呈現如圖1。



圖1 邏輯思考能力指標與選擇題各題型之對應

相較之下,閱讀測驗評量的思考能力涵蓋面明顯較廣,這或許與字彙題情 境、訊息較為有限,而綜合測驗、文意選塡、篇章結構則都為填空式命題方式, 資訊整合的範圍、靈活度頗受限制,而閱讀測驗試題則能靈活分列、比較、綜 合不同段落的訊息,更可彙整全篇文意,乃至跳脫全文之外,另外延伸、創造 一個新情境,要求考生加以判斷,命題的面向及靈活度遠非其他同屬選擇式試 題所能及。基於此點,很顯然地,若要評量較多的思考能力,不妨優先考量閱 讀測驗題。此外,歷年試題的分析也指出1-3指標「能分辨客觀事實與主觀意 見。」付之闕如,而前述模擬的第54題顯示,設計此類評量此能力的試題並非 不可能,值得命題者參酌。

在非選擇題部分,歷年學測與指考試題的題型中,各種型式(單句、連貫 式、克漏式)的翻譯題,由於各語意單位間的邏輯關係已經被中文提示句框定, 學生不需自行思考處理,無法評量學生的思考能力。句子合併與句子改寫多數 僅觸及兩句語意間的關係,所處理的訊息仍然相當有限,至多評量到1-1與1-2 兩項能力指標。作文部分,不管是簡函寫作、看圖作文、主題寫作或主題句寫 作,皆可涵蓋到多項的能力指標。以前述看圖作文爲例,可以評量到七項能力 指標,而分析資料也顯示,其他的作文試題,也都至少可以評量到1-1,1-2,2-1, 2-5 等四項思考能力。兩相比較,開放性較強的簡函寫作、主題寫作或主題句 寫作,由於沒有強制性,考生不一定會寫到、運用到某一些向度的思考能力, 所評量的思考能力面向反而可能較爲有限;相對地,看圖作文提供圖片,導引 學生朝一定的向度思考,可以比較出考生在特定向度思考能力的高低。

思考能力指標與非選擇式題型的對應關係,大致可以簡化呈現如圖2。



圖 2 邏輯思考能力指標與非選擇題各題型之對應

伍、結語

綜合上述研究發現與討論,本文一開始所提「評量思考能力的兩種可能 性」,答案至此便相當明確。雖然以往大學入學考試並未刻意評量思考能力,但 事實上,其學測與指考題型中,除了文法與翻譯題之外,各題型皆可以評量某 些面向的思考能力,呼應高中英文科課程綱要的思考能力指標。未來,大學入 學考試只要在命題設計時,除了原有的語言元素外,同時考量到不同面向思考 能力的評量即可,不需另闢專門的新單元評量思考能力。

不過,為了確實執行思考能力的評量,制訂適度的準則是必要的。以下僅 提供幾點原則性的淺見,供未來大學入學考試中心進一步研究參考。

45 🔳

一、各大題命題原則

- (一)各題型(除了翻譯題或文法題)測驗目標皆應納入思考能力。
- (二)各題型應該盡量涵蓋所可能評量到的各項能力指標。(例如:如上述資料 顯示,閱讀測驗題可能評量1-1, 1-2, 1-3, 2-1, 2-2, 2-3, 2-4七項能力指標, 整份試卷的閱讀測驗題便不宜僅評量到1-1, 1-2, 2-1三項。)

二、整卷組卷命題原則

- (一)考慮到少數題型(如綜合測驗的文法題、翻譯題)無法確實評量思考能力,未來初期,整份試卷宜至少有八成的試題能評量思考能力,之後再適時增加至九成。根據分析資料顯示,近年學測或指考試題,約有84%--90%的題目皆會觸及某一面向的思考能力,此八成的比例,應屬合理。
- (二)除了閱讀測驗與作文外,多數題型的思考能力評量侷限於1-1,1-2,2-1三 項能力指標,未來初期,整份試卷宜至少評量到八項思考能力指標中的五 項,之後再適時增加至六項或七項。根據分析資料顯示,近年每次學測或 指考整份試卷所觸及的思考能力多者有六項(如96學測),少者有四項(如 87年學測),每年整份試卷至少涵蓋五項不同的思考能力,應不難達成。 而涵蓋範圍的大或小,常取決於閱讀測驗題與英文作文題,值得命題者參 考。

本研究試著將以往的學測與指考試題加以分析,探索其中思考能力的成 分,並根據分析,提出各題型與能力指標的關係;然而,本研究雖投入相當多 的時間與人力,但仍屬拋磚引玉、探索性之研究,不盡周延之處必然不少,所 列的思考能力指標與題型對應關係表僅屬草案,尙祈先進指教、補正,更期待 未來有更多的研究人員投入探究,使得英文科思考能力評量的研究能更深入、 完備。

參考文獻

- 教育部(2008a)。普通高級中學英文科課程綱要補充說明。2011 年 3 月 10 日,取自 http://web.ylsh.ilc.edu.tw/course/files/990504/03.pdf。台北:作者。
- 教育部(2008b)。普通高級中學必修科目「英文」課程綱要。台北:作者。
- Bloom, P. and Keil, F. C. (2001). Thinking through language. *Mind & Language*, 16, (4), 351-167.
- Fisher, Robert. (2000). Thinking skills: Adding challenge to the curriculum. A guide for teachers of able children. Scottish Network for Able Pupils.
- Lin, Mei and Mackay, Cheryl. (2008). Thinking Through Modern Foreign Languages Chris Kington Publishing.
- Sapir, E. (1921). Culture, language, and personality. Berkeley: University of California Press.
- Whorf, B. L., & Carroll, J. B. (Eds.). (1956). Language, thought, and reality. Cambridge, MA: The MIT Press.

考試學刊 第九期 2011.06 Bulletin of Testing and Assessment

99 課綱與學科能力測驗自然考科相關議題之研究*

吳國良¹ 陳泰然²

大學入學考試中心1臺灣大學2

摘要

課程內容的調整會影響以課程為基礎的入學考試,本研究針對 99 課綱自然 領域的規範,研擬適用於學科能力測驗自然考科(簡稱學測自然考科)的測驗 範圍、試卷架構、題型與計分方式。經研究小組的方案設計、大規模問卷調查、 各型座談會與研究小組會議等步驟,建議 99 課綱學測自然考科的測驗範圍,仍 分爲第壹部分與第貳部分兩部分。但因 99 課綱高二必修的 8 學分比 95 課綱的 4~6 學分增加,建議第壹部分與第貳部分的占分比例為 62.5%: 37.5%。其次, 在計分方面,第壹部分每科 10 題,共 40 題,每題均計分,滿分 80 分。第貳部 分每科 7 題,共 28 題,答對 24 題,即得滿分 48 分。最後,在題型方面,第壹 部分採傳統題型,第貳部分則包括學科知識題與科學推理題,其中,第貳部分 科學推理題以 5 題爲原則,學科知識題以 2 題爲原則。科學推理題型是以測驗 考生數據分析、資料整理、推理論證相關的科學能力爲主,不涉及太艱深的學 科知識內容。

關鍵詞:學科能力測驗、自然考科、測驗範圍、99 課綱、95 課綱、必修學分、 題型、計分方式

吴國良,大學入學考試中心專門委員

陳泰然,臺灣大學大氣科學系教授

^{*}本文係根據大學入學考試中心民國 99 年計畫案:「新修訂課網(99 年)測驗內容與試題探究工作計畫一自然科」的研究報告改寫而成。研究小組成員包括:陳泰然(主持人)、張慶瑞、傅祖怡、劉緒宗、葉名倉、羅竹芳、林陳涌、李清勝、李通藝、程暐瀅、夏蕙蘭、吳國良。

A Study on Issues Related to the 2010 Edition of Curriculum Guidelines and GSAT-Natural Science

Kuo-liang Wu¹, Tai-lan Chen²

College Entrance Examination Center¹, National Taiwan University²

Abstract

This study focuses on issues related to the 2010 edition of Curriculum Guidelines and the GSAT-Natural Science. The 2010 edition of High School Curriculum Guidelines had been implemented in 2010, and the GSAT is one of the two major college admission examinations in Taiwan. The related issues include test material, test construction, item types and scoring of the GSAT-Natural science. The research group has gone through designing of testing material, large-scale questionnaire survey, and group discussions to reach a conclusion that under the Guidelines, the test construction of GSAT-Natural Science should be divided into two parts, with the test material of the first part including Basic Physics I, Basic Chemistry I, Basic Biology I theme 1-3, and Basic Earth Science I theme 1-5 and the second part including Basic Physics II A, Basic Chemistry II, Basic Biology I theme 4-6, and Basic Earth Science I theme 6-8. As the 2010 Guidelines requires 8 credits, more than 4-6 credits required by the 2006 Guidelines, it is suggested that the score distribution of these two parts be 62.5% and 37.5% separately, and the items of the first part include traditional types while the items of the second part include knowledge-based and science reasoning-based types. This new design of the GSAT-Natural science should be implemented in 2013.

Keywords: 2010-edition Curriculum Guidelines, GSAT-Natural science, test construction, test material

Kuo-liang Wu, Specialist, College Entrance Examination Center

Tai-lan Chen, Professor, Department of Atmospheric science, National Taiwan University

壹、前言

課程(curriculum)就其字義而言,是指跑馬道,可讓馬兒有所依循而不致 於迷失的一定路徑,衍生爲受教育所必經的歷程(賈馥茗等,2000),就學校的 系統來說,主要包括:課程內容(curriculum content)、教學(teaching)以及學 習(learning)。一般來說,課程的內容經由轉化之後便產生教科書(textbook) 或教材(teaching material),就台灣的情況而言,教科書的內容成了課程內容的 具體代表,是教師教學和學生學習的主要內容。

就過去將近三十年的歷史而言,我國的高中課程經過幾次的更迭與發展, 分別在民國 72 年、84 年、95 年與 99 年實施。各年代因時代潮流與實施方式的 不同,而有不同的名稱。例如:民國 72 年與 84 年實施的課程,稱之為「課程 標準」,亦即有成為一切準則的意涵。民國 95 年則改稱為「課程綱要」,即此乃 課程的大致規範,另 95 年所實施的課程,是為了配合九年一貫課程而暫時實 施,故稱為暫行綱要(教育部,2005);民國 99 年的課綱則是取代 95 年的暫行 綱要,於民國 99 年入學的高中生實施(教育部,2008)。上述不同年代所實施 的課程,本文統一以「實施年代 + 課綱」的方式表示,如:民國 99 年的課程 稱為 99 課綱。這些不同時期所發展的課程,內容會因應時代變遷而有所調整, 對學校的教學與學生的學習,有相當重要的影響。

課程內容的調整亦會影響以課程為基礎(curriculum-based)的入學考試。 理論上,以測驗內容對測驗進行分類,可以分成兩大類:一是性向測驗(aptitude test),這類的測驗是評量考生天生資質俱有的能力或技能,故比較屬於與學習 內容無關的測驗;另一類型是成就測驗(achievement test),是測驗考生學習後 所獲得的知識或技能,以目前台灣的入學考試而言,比較是屬於成就測驗的一 種。大學入學考試中心(以下簡稱大考中心)所設計的兩種大學入學考試,有 「學科能力測驗」(簡稱學測)與「指定科目考試」(簡稱指考),學測最主要是 做為推薦甄選的參考指標,目前包括繁星推薦與申請入學都屬於此一類別的入

51 🔳

學管道。當初設計學測的目的,是做為初步篩選學生之用,各校系須再利用第 二階段的指定項目甄試,選擇適合的學生,故學測以評量就讀大學的基本能力 為主,考生不分文理組,國文、英文、數學、社會與自然五科皆須報考,具通 識教育的意涵。考試的範圍是各科必修的內容,計分的方式採取較大量尺的級 分制。這種級分方式的設計,是要考生不必分分計較,即便考生錯了相當的題 數,依舊有獲得滿級分的機會。

就學測測驗設計的宗旨與測驗範圍來看自然考科,不管是配合過去的 72 課綱或是晚近的 84 與 95 課綱,皆是以高一(72 課綱)或高一與高二的必修(84 與 95 課綱)課程內容為其測驗範圍。同時,為配合 84 與 95 課綱中高二自然領 域採「四選一」的修習方式,自然考科的第貳部分採取了特殊的題型與計分方 式(吳國良等,2008)。不過,因 99 課綱與 95 課綱有所不同,如何設計出完善 的入學考試模式,是本研究嘗試達成的目標,主要的研究議題有二,分別是:

1. 探討 99 課綱下,學測自然考科的測驗範圍與試卷架構。

2. 探討 99 課綱下,學測自然考科的題型與計分方式。

貳、文獻探討

99 課綱自然領域包括物理、化學、生物與地球科學四科,以下就 **99** 課綱 自然領域修訂的精神、特色等,討論目前已有的相關研究或報告。

一、99 課綱自然領域修訂的精神

在探究 99 課綱學測自然考科的重要議題前,先行檢視 99 課綱的修訂精神, 可有助於釐清相關議題。就自然領域修訂精神而言,可簡化為三項,分別是通 識素養、延後分流與分版課程(李坤崇,2010)。若以高中課程乃全體國民應具 有的基本素養,則提高科學的通識素養正符合國際潮流與趨勢,檢視各國目前 自然領域與人文社會領域的修課學分,發現大部分國家皆採取 2:3 的比例。因 此,潘慧玲等(2005)建議,以目前台灣高中課程中人文社會領域必修學分數 為 24,建議自然領域必修學分數可為 16。此 2:3 的比例建議,正落實於 99 課綱的自然領域與人文社會領域。故就自然領域而言,與 95 課綱的 12-14 學分 相較,增加了 2-4 學分。此舉不僅拉近與人文社會領域學分的差距,也有助於 延後分流,對於社會組學生,在高二階段還能根據興趣與需求,選習 8 學分的 自然領域課程(教育部,2008)。但為賦予各校更大的開課空間,僅規範總必修 為 16 學分與各科至少修習 2 學分,高二學生可以有相當的修課彈性,例如文組 學生可能以生物與地科為優先考慮科目,二類組學生也可能只上物理與化學。 故設計自然考科的測驗範圍時,宜考量學生修課的彈性(吳國良等,2008)。

除上述通識素養與延後分流之外,99 課綱的另外一項特色是分版課程。此 種課程設計乃因應學生性向與生涯發展取向之差異,數學、英文、基礎物理三 科教材自高二起分為 A、B 兩版,提供不同深度、廣度與學習速度的課程,且 A 版教材包含於 B 版教材(教育部,2008)。這種分版課程的實施,除了有助 於各級學校選擇教材時,可考量學生的性向與發展之外,也讓文組學生若選習 物理時,可用難度較低的 A 版,不會太早放棄學習(楊思偉等,2005)。

二、99 課綱自然領域各科修訂的特色

99 課綱自然領域包括物理、化學、生物與地球科學四科,以下就四科課綱 修訂的特色進行介紹:

就物理而言,99 課綱物理科課程分為基礎物理與選修物理,基礎物理又分 為基礎物理一、基礎物理二A與基礎物理二B。基礎物理一是銜接國民中小學 九年一貫課程「自然及生活科技」學習領域教材,進一步介紹物理學基本知識, 此課程高一實施,為2學分的課程。基礎物理二A課程是屬於高二實施,為2 學分課程,主要是介紹力學的基本知識。基礎物理二B課程是於高二實施,為 4-6 學分之課程,主要是深入探索力學,並能瞭解物理學的基本精神與架構。 選修物理則是於高三實施,為一學年的課程,每學期4-5 學分。課程的目標為

增加學生對於物理瞭解的深度與廣度,訓練學生抽象思考、計算以及做實驗的 能力,以奠定學生研習下一階段物理的基礎(教育部,2008)。

就 99 課綱物理科的特色是著重通識教育的精神,以所謂綜論的方式介紹物 理學的知識,採取的是層次、架構與概念的漸進方式。簡單地說,就是「先見 林再見樹」,先讓學生對物理學的全局有初淺的認知,再進入細部的知識內容, 藉由說故事的方式,讓學生能夠了解與欣賞物理學的知識(高涌泉,2010)。這 種故事性質的敘述方式,希望培養學生的閱讀習慣,進而培養學生整合與敘事 方面的能力(高涌泉,2010)。這種著重通識、整合與敘事取向的理念,與國內 某些學者的想法不謀而合。例如,劉源俊(2009)曾提出,關於物理科的教材, 應分爲「基礎物理」與「進階物理」。「基礎物理」的授課對象是所有學生,其 宗旨是「文化素養」,主要內容包括:物理發展史(從天文革命講起);物理在 科學中的角色;物理與社會、科技、環境的關係;物理方法與科學精神;物理 之美。另外,「進階物理」的授課對象則是理科性向學生,其宗旨應是簡介完整 的物理架構並與現代科技接軌,主要內容包括從力學到量子物理的一般物理教 材,而且必須介紹一些歷史上關鍵的實驗與重要的推理過程。

此種重視文化素養與教材分級的想法,與 99 課綱物理領域部分,強調通 識、敘事與整合,以及物理教材分A、B版的設計,有相近的邏輯與理念。關 於評量部分,劉源俊(2009)的建議為:

(一)應重視問答題,減低選擇題的比重,因學生為練習速答選擇題,常背一

大堆公式,只注重代式子的運算,其實學不到觀念。

(二)「學科能力測驗」考「基礎物理」的範圍(素養)。

(三)「指定科目考試」考「進階物理」的範圍。

這種重視問答方式的評量方式,以及學測考基礎物理、指考包括進階物理 的想法,實與99課綱的設計理念十分貼近。

就化學而言,99 課綱化學科課程分為基礎化學、選修化學與選修化學實驗。與95 課綱相較,其修訂的方式採取「內容微調、結構大動」。修訂的三項

認知分別是:(1)自然科學內涵應含括自然界的課題與學科知識、科學方法與 科學思維模式、以及科學社群的互動。(2)科學或不代表真理,但是必須可以 理解。(3)學習者應發展出理性思辨與實事求是的核心能力與態度(陳竹亭, 2010)。就課程設計而言,基礎化學屬於必修課程,基礎化學(一)、(二)、(三) 為三學期課程,每學期2學分,內含化學實驗,其中基礎化學(一)、(二)是 以所有學生為修習對象而設計的課綱內容,基礎化學(三)則是以自然領域的 學生為考量而設計的課綱內容。選修化學為兩學期的課程,每學期3學分,選 修化學實驗同樣為兩學期的課程,每學期1學分,選修化學(含實驗)是以自 然領域的學生為考量而設計(陳竹亭,2010)。99 課綱化學科並未如物理課綱 一樣採取分版設計,其所強調的是知識與能力並重,且重視與科學社群的互動。 在評量內容部分,基礎化學(一)與基礎化學(二)屬於學測內容,基礎化學 (三)與選修化學則是屬於指考內容。

就生物而言,99課綱生物科課程分為基礎生物(1)、基礎生物(2)(應用 生物)與選修生物三個科目,其修訂的主要目標與理念是以現行「生物科課程 暫行綱要」為基礎,探漸進發展與微調修訂的方式進行,而不採革命性的改革 (鄭湧涇,2010)。就教育目標而言,重視認知、情意與技能三個領域。就考試 評量關係而言,生物科課綱小組認為,基礎生物(1)4學分的課程涵蓋生命的 特性、遺傳、演化與生物多樣性、植物與動物的構造與功能,以及生物與環境, 此六項主題涵蓋生物科的主要主題,構成完整的基本生物科學素養,因此,建 議此4學分列為全體高中學生必修,並納入學測自然考科的測驗範圍(鄭湧涇, 2010)。至於,基礎生物(2)(應用生物)則是以議題為導向,建議不納入學測 與指考的測驗範圍。而選修生物則是納入指考的測驗範圍,即指考生物考科的 測驗範圍包括基礎生物(1)與選修生物。

最後,就地球科學而言,99課綱地球科學的課程分為:必修基礎地球科學 與選修基礎地球科學,其中必修基礎地球科學為4學分課程,可安排二學期, 每學期2學分。第一學期規劃第一個主題至第五個主題;第二學期規劃第六個

主題至第八個主題。選修基礎地球科學則是2學分的課程,於高二實施。「基礎 地球科學」的課程目標旨在培養具備「地球科學」基本素養,包括知識、思維、 技能、態度的現代公民(教育部,2008)。

綜合上述,針對本研究之重要議題,即研擬 99 課綱學測自然考科的測驗範 圍與試卷架構,以及題型與計分方式,提供方案設計的參考。

- (一)在95課綱下,自然領域高一與高二的必修學分是8+4~6。在99課綱下,自然領域高一與高二的必修學分則是8+8。故就自然領域而言,99 課綱較95課綱增加2~4學分。
- (二)99 課綱與95 課綱均有相同的規定,即基礎物理、基礎化學、基礎生物 與基礎地球科學每科至少修習2學分。
- (三)自然領域四個學科中,有三個學科的學科專家或課綱修訂召集人,對於 學測的測驗範圍有所建議:物理科建議考基礎物理、化學科建議包括基礎化學(一)與基礎化學(二)、生物科則建議考基礎生物(1)。

三、學測自然考科的現況

自然領域的課程設計,將直接影響學測自然考科的測驗範圍、試卷架構、 測驗題型,乃至於計分方式。就 95 課綱而言,學測自然考科的試題分為兩部分, 其試卷架構、測驗內容、題型題數與計分方式如表 1 所示(大學入學考試中心, 2005)。

表1 95 課綱學測自然考科的試卷架構

試卷架構	測驗內容	題型與題數	計分方式
第壹部分	高一爲主	各科 12 題,傳統題型	全部計分
第貳部分	偏重高二	各科5題,4題科學推理題型, 1題學科知識題型	答對一定比例即得滿分(20題 答對16題即得滿分)

從表1來看,自然考科的試卷分為第壹部分與第貳部分,第壹部分以評量 高一課程內容為主,第貳部分則偏重高二課程內容。第壹部分為傳統題型,各 科各12題;第貳部分為科學推理題型,測驗考生數據分析、資料整理以及推理 論證等能力,此題型不涉及太艱深的學科知識內容,且在計分上是採答對一定 比例即得滿分的方式。此種題型與計分方式的設計,主要是考量學生在高二時 乃是採「四選一」的修習方式,即包括物理、化學、生物、地球與環境四科, 並未規定要修習哪一科目,但每學期至少選習2至3學分。為配合此種課程設 計,故以物理、化學、生物與地科四個科目共通的「科學方法」,作為命題的依 歸。此種命題方式從91年學測開始實施,已有多年的施測經驗(吳國良等, 2008)。

綜合上述文獻探討的內容,對 99 課綱學測自然考科的測驗範圍、試卷架 構、題型與計分方式的探討,已具初步的方向。問題的核心在於測驗範圍與試 卷架構,即 99 課綱學測自然考科的試卷是否依循 95 課綱的設計,分爲第壹部 分與第貳部分,若分爲兩部分,則其涵蓋的範圍爲何?以下將藉由不同的研究 方法,逐一探討。

參、研究的方法與步驟

本研究採取兩個階段進行,第一階段研擬與確定 99 課綱自然考科的測驗範 圍與試卷架構,進行的時間為 99 年 1 月至 4 月;第二階段則研擬測驗題型與計 分方式,進行的時間為 99 年 5 月至 6 月,研究的步驟如下:

一、課綱研究與問卷設計

首先,研究小組針對 99 課綱進行研究,並對學測自然考科測驗範圍可能採 取的方案,就各個面向進行考量。評估方案的優缺點,討論可能遭遇的問題, 並歸納出合適的方案。其次,以此方案設計問卷,內容著重於 99 課綱的設計精

神與理念,各學科課程設計的差異、學測自然考科的基本理念、以及大學選才、 高中排課的行政等。

二、召開小型座談會

初擬問卷後,邀請問卷設計專家,高中教務主任與自然領域的教師代表約 15~20人,進行小型座談會,並對問卷的內容提出修改建議,作為大規模問卷 的基礎。

三、大規模問卷調查

以修正過後的問卷,採普查方式進行大規模調查,調查對象包括高中自然 類科的教師、各校的教務主任以及大學所有學系的系主任。

四、召開中型座談會

綜合問卷的結果,研究小組對 99 課綱自然考科的測驗範圍與試卷架構提出 可行方案,以兩案並呈方式,進行中型座談會。會議廣邀各界人士,包括:教 育部代表、教師團體、家長團體、高中校長代表、自然領域各學科 99 課綱召集 人、學科教授等,就研究小組提出的方案,進行意見交換。座談會後,研究小 組整理各界對不同方案的意見,提交大考中心的行政會議進行討論與決議。

五、題型與計分方式研究

待測驗範圍與試卷架構確定後,研究小組針對題型與計分方式進行研究, 研究的議題包括:是否沿用以往的科學推理題型、科學推理題所占的比例、以 及是否仍採答對一定比例即得滿分的方式等。

肆、結果與討論

依本研究步驟,以下分述結果與討論。

一、課綱研究與問卷設計

首先,研究小組針對 99 課綱進行研究,並對學測自然考科的測驗範圍,可 能採取的方案,就各個面向進行考量,歸納出四個具體的方案,說明如下:

(一)方案一

参照 95 課綱的作法,將自然考科的試卷分為兩部分,第壹部分評量內容為 每學科 2 學分,包括:基礎物理一、基礎化學(一)、基礎生物(1)上半部、 基礎地科(主題 1-5);第貳部分則各學科只包含 2 學分,包括:基礎物理二 A、 基礎化學(二)、基礎生物(1)下半部、基礎地科(主題 6-8)。第壹部分與第 貳部分的四個科目加起來總共 16 學分,符合總綱中高一與高二合計需修習 16 學分的要求(見表 2)。

表2 方案一

		第壹部分測驗範圍	學分	第貳部分測驗範圍	學分
		基礎物理一	2	基礎物理二 A	2
	自然考科	基礎化學(一)	2	基礎化學(二)	2
	基礎生物(1)上半部	2	基礎生物(1)下半部	2	
	*基礎地科(主題 1-5)	2	*基礎地科(主題 6-8)	2	

註:*基礎地科的主題 1-5 與 6-8 的切割,爲課綱之規定。

(二)方案二

自然考科的測驗範圍不再分第壹部分與第貳部分,物理、化學、生物與地 科四科的測驗範圍都包括4學分,合計是16學分,包括:基礎物理一、基礎物 理二A;基礎化學(一)、基礎化學(二);基礎生物(1)全部範圍、基礎地科 (見表3)。

表3 方案二

	測驗範圍	學分
	基礎物理一、基礎物理二 A	$2\!+\!2$
自然考科	基礎化學(一)、基礎化學(二)	$2\!+\!2$
	基礎生物(1)	4
	基礎地科	4

(三) 方案三

自然考科的測驗範圍包括各領域2學分,合計是8學分(見表4)。

表4 方案三

	測驗範圍	學分
	基礎物理一	2
自然考科	基礎化學(一)	2
	基礎生物(1)	2
	基礎地科	2

(四) 方案四

自然考科的測驗範圍包括第壹部分各學科的2學分,合計8學分,第貳部 分由考生任選2科,且只能選2科,合計是8學分(見表5)。

表5 方案四

	第壹部分測驗範圍	學分	第貳部分測驗範圍	學分
自然考科	基礎物理一	2		
	基礎化學(一)	2		4
	基礎生物(1)	2	四科仕選兩科	4
	基礎地科	2		

研究小組就上述四種方案的優缺點,可能遭遇的問題進行討論,重點包括: 自然領域四科之間須採一致的作法,若某一科目作法有所不同,則不易對外說 明;各分科間課程設計理念並不相同:基礎生物(1)是採4學分完整成套的設

計;基礎地科採4學分設計,但其內容可分成兩部分;基礎物理與基礎化學則 採2學分的設計方式;基於學測的測驗範圍包括必修課程內容,而方案三只考 8學分,並不符合教育部要求99課綱自然領域高一、高二必修16學分的規定; 另外,就本中心實際可操作性,如方案四第貳部分採選科考的方式,各科目之 間難度不一致的問題,並不容易解決。

綜合上述的論點與考量,研究小組初步認為方案一與方案二是較可行的方 案,為了解高中端與大學端對此二種方案的想法與支持度,研究小組設計問卷 進行調查。問卷的內容,著重於99課綱的設計精神與理念,各學科課程設計的 差異、學測自然考科的基本理念、以及大學選才、高中排課的行政等。問卷的 初稿即為小型座談會討論的基礎。

二、小型座談會

問卷初稿完成後,邀請問卷設計專家,高中教務主任與高中自然領域的代表,於99年2月27日於本中心進行小型座談會,並對問卷的內容提出修改建議,作為大規模問卷的基礎。修改完成後的問卷定稿內容,包括問卷的說明, 基本資料與各項有關測驗設計方案的相關問題等。就大學部分,基本資料包括 學系使用學測自然考科作為入學管道以及使用的方式;就高中部分,再分為教 務主任與自然領域學科教師的版本,基本資料是包括任教學校名稱、任教學校 的屬性與任教的科目(定稿之問卷內容如附錄一)。

三、大規模問卷調查

上述問卷進行大規模的調查,調查的對象包括高中自然類科的教師、各校 的教務主任以及各大學校系的系主任。問卷調查的結果,說明如下:

(一)大學部分

1. 採普查方式,以校為寄送單位,共寄出 65 所大學,全部的學系數為 1426
 個,回收 1217 份,回收率為 85%。

- 回收的 1217 個學系中,採用學測自然考科作為入學標準者有 476 個學 系(39%),未採用學測自然考科作為入學標準者有 672 個學系(55%),有 69 個學系(7%)未對此題回覆。對於方案一與方案二的選擇,有 208 個學 系(17%)選擇方案一,208 個學系(17%) 選擇方案二,523 個學系(43%) 則是認為方案一或方案二差異不大,任一方案均可,由大考中心決定。
- 對於學系採方案一的原因,前三名依序為:降低考生的學習負擔、符合 學科知識結構的完整性、減輕考生的應考壓力。而採方案二的原因,前 三名依序為:符合學科知識結構的完整性、增加命題的廣度、增進學生 學習的機會。
- 4. 就回覆採用學測自然考科作為入學標準的 476 個學系統計,有 117 個學 系(25%)贊成採方案一,115 個學系(24%)贊成採方案二,225 個學系(47%) 則認為採方案一或方案二均可。至於認為採取方案一或方案二皆可的主 要原因,則與其他學系的選答結果類似,僅排序略作更動。
- 5. 從大學學系問卷結果發現:大部分的學系認為方案一與方案二差別不 大,擇一實施均可;選方案一或方案二的學系數目大致相當,無論是就 全部回覆的學系,或是就有採用學測自然考科成績的學系,皆得到類似 的結果,表示對於方案的選擇並未因是否採用學測自然考科成績與否而 有明顯不同,整理如表 6。

校系類別	贊成方案一 (%)	贊成方案二 (%)	方案一或方案二均可 (%)
全部回覆學系(1217)	17	17	43
採用學測自然考科成績	25	24	47
學系(476)	20	27	1 ד

表6 大學學系贊成方案的統計

(二)高中部分

- 1. 採普查方式,以校為寄送單位,共寄出317所學校,每校5份,分成教務主任1份與自然領域教師4份。
- 2. 教務主任回收 164 份,回收率為 52%。其中,有 106 份(65%)贊成採方 案一,有 38 份(23%)贊成採方案二,有 18 份(11%)認為方案一與方案二 均可。至於採方案一的原因,前三名依序如下:降低考生的學習負擔、 符合學測作為進入大學基本門檻的目標、減輕考生的應考壓力;採方案 二的原因,前三名依序為:符合學科知識的完整性、符合 99 課綱延後 分流的精神、學習內容較完整符合大學選才的需求。
- 3. 高中自然領域教師部分,寄出 1268 份,回收 538 份,回收率為 42%。 其中,有 186 份(35%)贊成採方案一,有 271 份(50%)贊成採方案二,有 75 份(14%)認為採方案一或方案二均可。至於採方案一的原因,前三名 依序為:降低考生的學習負擔、減輕考生的應考壓力、符合學測作為進 入大學基本門檻的目標;採方案二的原因,前三名依序為:符合學科知 識結構的完整性、符合 99 課綱延後分流的精神、增進學生學習的機會。
- 4. 高中部分的初步結論:高中教務主任較贊成方案一,自然領域的教師則 較贊成方案二,結果整理如表 7。不論是教務主任或是自然領域教師選 擇方案一或方案二的主要原因,大致呈現相類似的結果。

教師別	贊成方案一 (%)	贊成方案二 (%)	方案一或方案二均可 (%)
教務主任(164)	65	23	11
自然領域(538)	35	50	14

表 7 高中贊成方案的統計



四、中型座談會

綜合問卷的結果,研究小組對 99 課綱自然考科的測驗範圍,進行中型座談 會,邀請各界人士參與,包括:教育部代表、高中校長、家長團體、教師團體、 99 課綱自然領域課綱召集人等,於 99 年 5 月 3 日在本中心,就問卷結果交換 意見。會議中,大部分的高中校長與全國教師會代表發言,均贊成方案一,認 為方案一較具彈性,因就不同類組的學生而言,可有不同的選課方式,一類組 的學生可以修習物理、化學、生物與地科,二類組的學生可以修習物理與化學, 三類組的學生可修習物理、化學與生物,而方案二的內容,各組的學生必須修 物理、化學、生物與地科四個學科,否則無法因應考試的內容,故就目前高中 的師資或排課,並無法配合此種改變。然會中也有反對方案一的聲音,因方案 一須將基礎生物(1)課程分成上、下兩部分,生物學者認為基礎生物(1)涵 蓋生物學的六大主題,若要從任何一個內容進行切割,是不可能的。而且,考 試所面對的問題,應以測驗技術去克服,或是四個科目可採不同的作法,較具 彈性,不應對課綱內容進行任何調整。

五、題型與計分方式研究

經過問卷調查結果發現,大學校系對 99 課綱學測自然考科的測驗範圍,贊 成採方案一或方案二,無明顯的差別;高中端則是行政系統與自然領域教師呈 現不同的想法。再加上中型座談會中,大部分的與會人員皆認為方案一較具彈 性,並可與高中教學現場配合,故研究小組建議 99 課綱學測自然考科的測驗範 圍採方案一。即 99 課綱學測自然考科的試卷分為兩部分,第壹部分評量內容為 每學科 2 學分,包括:基礎物理一、基礎化學(一)、基礎生物(1)上半部、 基礎地科(主題 1-5);第貳部分各領域亦包含 2 學分,包括:基礎物理二 A、 基礎化學(二)、基礎生物(1)下半部、基礎地科(主題 6-8)。雖然研究小組 建議,99 課綱學測自然考科的試卷採兩部分的設計,然座談會中,仍有學者認 為基礎生物(1)的4 學分為一完整課程,不宜進行切割;再加上,本中心並不

屬於行政機關,不應對課程進行解釋或調整,故此項工作是由教育部委託課程 發展工作圈進行,所得到的結論是:99課綱學測自然考科,就生物而言,第壹 部分的測驗範圍是基礎生物(1)主題壹、貳、參;第貳部分的測驗範圍是基礎 生物(1)主題肆、伍、陸(教育部,2010)。

此外, 就方案一而言, 有幾項待討論事項, 分別是: 第壹部分與第貳部分 的占分比例為何、第貳部分答對多少比例即可得滿分、第貳部分的題型等。針 對第貳部份的占分, 研究小組提出甲、乙、丙、丁四個方案, 分析如表 8。表 8 中, 95 課綱自然考科第貳部分的占分比例是 25%, 而甲、乙、丙、丁四案的占 分比例分別是 30%、37.5%、45%、50%, 研究小組經討論後, 認爲乙案較爲合 理, 因爲依據 95 課綱, 高一與高二自然領域選習的學分比爲 8:4~6, 自然考 科第壹部分與第貳部分的配分比爲 96:32, 即 3:1。若依 99 課綱, 高一與高 二的選習學分比爲 8:8, 高二部分必修學分數增加, 則第貳部分的占分依比例 也應增加。計算如下:

95課綱高二自然領域必修學分/95課綱(高一+高二)自然領域必修學分 99課綱高二自然領域必修學分/99課綱(高一+高二)自然領域必修學分 = 95課綱自然考科第貳部分占分比例 99課綱自然考科第貳部分占分比例

狀況一:學生在 95 課綱高二自然領域選習 4 學分,則高一與高二自然領域必修 學分為 12,若 99 課綱自然領域第貳部分的占分比例為 x₁,則

$$\frac{\frac{4}{12}}{\frac{8}{16}} = \frac{25\%}{x_1} \qquad \chi_1 = 37.5\%$$

狀況二:學生在95課綱高二自然領域選習6學分,則高一與高二自然領域必修 學分為14,若99課綱自然領域第貳部分的占分比例為x,,則

$$\frac{\frac{6}{14}}{\frac{8}{16}} = \frac{25\%}{x_2} \qquad \chi_2 = 30\%$$

故 99 課綱學測自然考科第貳部分的占分比例,30%與 37.5%皆可接受。因 考量 99 課綱具有延後分流的精神,故研究小組建議採取乙方案:37.5%。若依 此比例與表 8 中乙案的內容,研究小組建議:99 課綱學測自然考科第壹部分每 科 10 題,共40 題,每題均計分,共80分;第貳部分每科 7 題,共28 題,答 對 24 題即得滿分 48分,兩部分的占分比為 62.5%:37.5%。第壹部分採傳統題 型,第貳部分的試題,可採學科知識題型與評量科學方法的科學推理題型。不 過,研究小組就第貳部分中知識題與推理題的題數,並未達成共識。

為解決第貳部分學科知識題型與科學推理題的比例,本中心於99年10月 21日進行「99課綱學測自然考科試卷架構座談會」,由台灣大學大氣科學系陳 泰然教授擔任主持人,邀請人員包括:教育部官員、大學教授、高中校長、學 科中心教師代表出席,本中心副主任、2位顧問、四位自然科學科研究員列席 參與。會中達成共識為: 99課綱學測自然考科第貳部分每科7題,學科知識 題以2題為原則、科學推理題以5題為原則。

方式	第壹部分	第貳部分	第貳部分 占分比例	備註
95 課綱	48題96分	20題滿分 32分	25%	第貳部分答對 <u>16</u> (80%) 即得滿分 32 分
99 課綱 甲案	44題88分	24.題滿分 40分	30%	第貳部分答對 <u>20</u> (83%) 即得滿分 40 分
95 課綱 乙案	40題80分	28題滿分 48分	37.5%	第貳部分答對 24 28(86%) 即得滿分 48 分
95 課綱 丙案	36題72分	32.題滿分 56 分	45%	第貳部分答對 <u>28</u> 32(87.5%) 即得滿分 56 分
95 課綱 丁案	32題64分	36題滿分 64分	50%	第貳部分答對 <u>32</u> 36(88.8%) 即得滿分 64 分

表8 學測自然考科第壹部分與第貳部分的占分比例

就整個研究的討論過程而言,充滿了各種不同的意見與聲音,期間若不計 外部會議,大考中心共舉行了四場座談會。包括:計畫開始執行前的議題溝通 (98年11月)、問卷發出前的專家審查會議(99年2月)、問卷結果討論會議 (99年5月)、以及試卷架構會議(99年10月),此外,尙有教育部的五方座 談會(99年6月)與課程發展會議(99年7月)等。在此其間,對99課綱學 測自然考科相關議題所投注的人力、物力與各方的資源,確實相當可觀。這些 討論的議題,包括第壹部分與第貳部分的測驗範圍,尤其是考量生物課程的完 整性;第壹部分與第貳部分的配分比例;以及第貳部分的題型,包括學科知識 題與科學推理題的題數等,都是費盡心思並且不斷來回討論,所達成的一個折 衷方案。目前這個方案,各界仍舊有不同的聲音,但這是經過多項程序、多位 學者專家、高中校長與教師討論而得,若需調整,直經更嚴謹的程序。

課程的分配就如同權力的分配,考試只是課程實施過程中,最末端的部分。 若欲藉由考試或是入學考試來解決課程的爭端,無異於緣木求魚、治絲益棼。 況且,大考中心本身是一個負責試務的獨立財團法人機構,實無任何法律上的 位階,無法解決屬於課程訂定所衍生的相關問題。在課程訂定的過程中,實宜 考量對各種層面的影響,包括:學分重新分配後的師資結構、入學考試內容以 及國內高中生基本學科素養等,這些問題實不宜由單一委員會決定,應由專業 機構進行長期的研究與探討,才是國家教育的長遠之道。

伍、結論與建議

一、結論部分

關於 99 課綱自然考科測驗範圍、試卷架構、題型與計分方式,經研究小組的研議、大規模問卷調查、各型座談會的討論、課務發展工作圈的協助調查之後,獲得的結論如下:

(一)99課綱學測自然考科的測驗範圍如表9所示,自然考科的試卷分為第壹 部分與第貳部分兩部分。

表9 99 課綱學測自然考科的測驗範圍

	【第壹部分】測驗範圍	學分	【第貳部分】測驗範圍	學分
白佛	基礎物理一	2	基礎物理二 A	2
目然 考科	基礎化學(一)	2	基礎化學(二)	2
	基礎生物(1)(主題 壹~參)	2	基礎生物(1)(主題 肆~陸)	2
	基礎地科(主題 1~5)	2	基礎地科(主題 6~8)	2

其中,第壹部分的測驗範圍包括:基礎物理一;基礎化學(一);基礎生物(1)主題壹、貳、參的內容;與基礎地科主題 1~5 的內容。第 貳部分的測驗範圍包括:基礎物理二A;基礎化學(二);基礎生物(1) 主題肆、伍、陸的內容;與基礎地科主題 6~8 的內容。其中,基礎生物 (1)的測驗範圍是根據教育部的普通高級中學課務發展工作圈所作的建 議而決定(教育部,2010),主要考量考生第壹部分的 8 學分是全部修習, 而第貳部分的 8 學分,可能因類組不同,而有不同的選習方式。另外, 99 課綱高二必修的 8 學分比 95 課綱的 4~6 學分增加,參考 95 課綱學測 自然考科第貳部分的占分比例為 25%,建議 99 課綱學測自然考科第壹 部分與第貳部分的占分比例為 62.5%: 37.5%。

- (二)第壹部分每科 10題,共40題,每題均計分,滿分 80分。第貳部分每科
 7題,共28題,答對24題,即得滿分48分。兩部分加起來總分是128
 分,與95課綱學測自然考科的總分及總題數相同。
- (三)第壹部分採傳統題型,第貳部分則包括學科知識題與科學推理題,其中, 第貳部分科學推理題以5題為原則,學科知識題以2題為原則。科學推 理題型是以測驗考生數據分析、資料整理、推理論證相關的科學能力為 主,不涉及太艱深的學科知識內容。

二、建議部分

就課程設計而言,評量實屬於課程設計的一環,尤其是入學考試的內容, 因牽涉層面很廣,宜在課程設計之初,即進行通盤考量。另外,此方案已於大 考中心的網站公告,希望各界能夠儘早有所準備。就大考中心而言,針對 99 課綱學測自然考科的題庫建置,也已著手努力進行中。

參考文獻

大學入學考試中心(2005)。認識學科能力測驗8。台北市:作者。

- 李坤崇(2010)。高中課程 99 課綱與 95 暫綱之分析。**教育資料與研究,92**,1-24。台 北市:國立編譯館。
- 吳國良、蔡尙芳、邱美虹(2008)。學科能力測驗自然考科的發展歷史與未來展望。考 試學刊,4,27-57。
- 高涌泉(2010)。從高中物理課綱看考試的方法與內容。科學素養、科學課程與評量研 討會會議手冊。台北市:國立台灣師範大學科學研究所。
- 教育部(2005)。普通高級中學課程暫行綱要。台北市:作者。
- 教育部(2008)。普通高級中學課程綱要。台北市:作者。
- 教育部 (2010)。台中 (三) 字第 0990144290 號文。台北市:作者。
- 陳竹亭(2010)。普通高級中學化學科 99 課綱制定理念、特色與教科書編審用考的教 育功能。考試學刊,8,137-146。
- 賈馥茗、伍振、楊深坑、黃發策、謝福生(2000)。教育大辭書。台北市:國立編譯館。 楊思偉、李咏吟、高新建、黃淑馨、李振賢、陳盛賢(2005)。普通高中課程分級之研
 - **究**。教育部委託專案研究報告。台北市:國立台灣師範大學教育學系。
- 潘慧玲、王如哲、葉興華、張炳煌、吳武雄、吳清鏞(2005)。**普通高中課程必選修之** 研究。教育部委託專案研究報告。台北市:教育部。
- 劉源俊(2009)。論高中的物理教育。**物理雙月刊,31**,384-389。台北市:中華民國物 理學會。
- 鄭湧涇(2010)。由高中生物科課程綱要的修訂談生物科的教學與評量。科學素養、科 學課程與評量研討會會議手冊,國立台灣師範大學科學教育研究所。

附錄一

大學入學考試中心

普通高中 99 課綱學科能力測驗自然考科

測驗範圍問卷調查(高中教師版)

主任/老師您好:

99 課綱即將於九十九學年度正式實施,此課綱與目前實施的 95 課綱 有所不同。就自然領域而言,其中一項重要的不同是必修學分數的增加, 99 課綱必修學分比 95 課綱增加 2~4 學分,99 課綱自然領域的課程設計請 參見附件。

為因應新課綱自然領域的課程設計,學科能力測驗自然考科(簡稱學 測自然考科)的測驗範圍必須重新思考。學測的設計目的旨在評量考生是 否具有接受大學教育的基本學科能力,作為大學校系初步篩選學生的門 檻。本研究小組針對 99 課綱學測自然考科測驗範圍,設計了兩種方案, 詳細內容與特色請參見附件。本問卷針對這兩種方案,就教於各位先進。 懇請撥冗填答,並交回主任再以所附的回郵信封寄回大考中心,謝謝。

新修訂課綱(99年)測驗內容與試題探究工作計畫-自然科研究小組 敬上

聯絡人: 吳國良先生 電 話: 02-23661416#208

第一部分、基本資料

1. 校名:_____

2. 您所任教的學校是屬於:

	□ A.國立	□ B.市立	□ C.縣立	□ D.私立
3.	您所任教的學科	:		
	□ A. 國文	□ B. 英文	□ C. 數學	
	□ D. 歷史	□ E. 地理	□ F.公民與社會	
	🗌 G. 物理	□ H.化學	□ I. 生物	□ J. 地球科學
	□ K. 其他			
第二部分、問卷內容

- 1. 您認為 99 課綱學測自然考科的測驗範圍,下列何者較合理?
 - □ A.方案一
 - □ B.方案二
 - □ C.方案一與方案二差異不大,任一方案均可,由大考中心決定 (以下問題可免答)
 - □ D.其他,請說明
- 2. 您贊成 99 課綱學測自然考科的測驗範圍,採方案一或方案二的原因為 何?(可多選,至多選三項)
 - □ A.符合 99 課網延後分流的精神
 - □ B.增進學生學習的機會
 - □ C.降低考生的學習負擔
 - □ D.減輕考生的應考壓力
 - □ E.符合學科知識結構的完整性
 - □ F.增加命題廣度
 - □ G.學校排課較有彈性與師資較易配合
 - H.學校可增聘相關的師資人員
 - □ I.學校可減少相關的師資人員
 - □ J.符合學測作為進入大學基本門檻的目標
 - □ K.學習內容較完整符合大學選才的需求
 - □ L.吸引文組學生選讀理組科系
 - □ M.其他,請說明_____

附件

1. 研究小組設計的方案一與方案二

方案一的特色:

參照目前 95 課綱的作法,將自然考科的試卷分兩部分,第壹部分評量 內容為每學科2學分,考生每題均須作答與計分;第貳部分則各領域只包含 2學分,四個科目加起來總共8學分,即符合總綱中高一與高二至少皆需修 習8學分的要求。

	【第壹部分】測驗範圍	學分	【第貳部分】測驗範圍	學分
自然考科	基礎物理一	2	基礎物理二A	2
	基礎化學(一)	2	基礎化學(二)	2
	基礎生物(1)上半部	2	基礎生物(1)下半部	2
	*基礎地科(主題 1-5)	2	*基礎地科(主題 6-8))	2

方案一

註:*基礎地科的主題1-5與6-8的切割,為課網之規定

方案二的特色:

自然考科不再分第壹部分與第貳部分,物理、化學、生物與地科四科都 考4學分。

	測驗範圍	學分
	基礎物理一、基礎物理二 A	2 + 2
自然考科	基礎化學(一)、基礎化學(二)	2 + 2
	基礎生物(1)	4
	基礎地科	4

方案二

2. 目前學測自然考科的測驗範圍

	【第壹部分】測驗範圍	學分	【第貳部分】測驗範圍	*學分
	基礎物理	2	物理	6
自然考科	基礎化學	2	化學	6
	基礎生物	2	生物	4
	基礎地科		地球與環境	4

95 課綱自然考科測驗範圍

註:學生在高二的四個科目中,至少須選習一個科目,即至少須選習 4~6 學分

- (1)目前學測自然考科的測驗範圍分為第壹部分與第貳部分,第壹部分以 評量高一課程內容為主,第貳部分則是偏重高二課程內容。
- (2)第壹部分各科12題,屬傳統題型;第貳部分有4題是以測驗考生數據 分析、資料整理、推理論證相關的科學推理題型,此題型不涉及太艱深 的學科知識內容,有1題是學科知識題。
- (3)第貳部分的計分採答對一定比例即得滿分的方式,此項計分方式的設計,是考量學生目前在高二課程是採「四選一」的修習方式,即二年級包括物理、化學、生物及地球與環境四科,考生必須至少選擇一科修習。因為大部分的考生在高二並未完整學習四個科目,考量考生學習的壓力,故採答對一定比例即得滿分的方式。

3.99 課綱自然領域的課程設計

根據教育部於97年1月所頒布的《普通高級中學課程綱要》(以下簡稱99課綱),自然領域的課程設計如表所示:

領域	科目名稱	修課時間	學分	備註
	*基礎物理一	高一/一學期	兩學分	基礎物理
	*基礎物理二A	高二/一學期	兩學分	至少修習
物理	*基礎物理二B	高二/一學年	四至六學分	二學分
	選修科目「物理」	高三/一學年	每學期四至五	
			學分	
	*基礎化學(一)	高一/一學期	兩學分	基礎化學
7. 與	*基礎化學(二)	高二/一學期	至少兩學分	至少修習
化学	*基礎化學(三)	高二/一學期	至少兩學分	二學分
	選修科目「化學」	高三/二學期	每學期四學分	
	*基礎生物(1)	高一或高二/二學期	四學分	基礎生物
4	*基礎生物(2)	高二/一學期或一學年	二學分	至少修習
生初	(應用生物)			二學分
	選修科目「生物」	高三/一學年	八學分	
	*基礎地球科學	高一或高二/二學期	四學分	基礎地科
地球	選修科目「基礎地球	高二/一學期	二學分	至少修習
科學	科學」			二學分
	選修科目「地球科學」	高三/二學期	四學分	

99 課綱自然領域課程設計

註:有"*"者表必修科目。

由表中可知,自然領域的課程設計,有兩個重點,分別是:

- 自然領域包括基礎物理、基礎化學、基礎生物、基礎地球科學等四科,每一 科目至少修習2學分。
- 自然領域高一與高二,各學期(上學期與下學期)必修4學分,總共必修學 分數為16。

考試學刊 第九期 2011.06 Bulletin of Testing and Assessment

地球科學動畫試題的設計與應用

吳皇慶 張俊彦

臺灣師範大學

摘要

本實驗開發了一套結合動畫與測驗的動畫試題,試題中運用動畫來呈現地 球科學領域當中包含天文、大氣、海洋與地質等內容。本研究的目的在於設計 出一套信度與效度皆符合標準的動畫試題,並比較學生在動畫試題與傳統試題 的答題表現與答題態度。研究發現,動畫試題具有足夠的信度與效度,且發現, 先備知識較低的學生在動畫試題的表現優於傳統的圖文試題,而其餘的學生則 沒有顯著的差別。答題態度方面,學生普遍對於動畫試題感到有趣並樂於接受 此一評量方式。

關鍵詞:動畫、電腦化測驗、成就、態度



吳皇慶,臺灣師範大學地球科學所博士班學生

張俊彥,臺灣師範大學研究講座教授,科學教育中心主任

The Design of an Animation-Based Test System in the Area of Earth Sciences

Huang-Ching Wu, Chun-Yen Chang

National Taiwan Normal University

Abstract

This study developed a technology-enhanced testing device, namely, the animation-based test (ABT) in the area of earth sciences ranging from astronomy, meteorology, oceanography, to geology. Both the students' achievements of and their attitudes toward ABT were compared with those of and those toward the traditional graphic-based test. The purposes of this study were fourfold: (1) to examine the validity and the reliability of ABT, (2) to detect the difference of ABT and GBT regarding students' achievements, (3) to investigate the impact of ABT versus GBT on students' achievements with different levels of prior knowledge, and (4) to explore the ABT participants' attitudes toward ABT. The results of the study indicated that ABT was a valid and reliable way of testing. It was found that students with low prior knowledge level performed better in ABT while students with high prior knowledge level performed equally well in both tests. Finally, the participants who took ABT were satisfied with and held positive attitudes toward ABT.

Keywords: animation, computerized assessment, achievement, attitude

Huang-Ching Wu, Doctoral student, Department of Earth Sciences of NTNU Chun-Yen Chang, Research Chair Professor, NTNU; Director, Science Education Center of NTNU

壹、前言

一直以來,研究者致力於設計出能夠準確測量學生科學成就的試題,並藉 此判斷學生是否達成學習目標。傳統的試題多半採用紙筆測驗的方式,這是學 生相當熟悉的測驗方式,並且容易執行,也方便批閱。儘管如此,紙筆測驗在 呈現試題上,有著部分的限制,這些限制包括:(一)呈現抽象科學概念時,即 使圖文並茂,仍有可能讓測驗者無法清楚了解題意。舉例來說,許多科學概念 涵蓋大尺度的現象,如地球科學領域中所提到的板塊運動或者地球、月球與太 陽間的運行關係;也有科學概念包含微尺度的現象,如化學領域中所談到的原 子分子理論。上述這些概念如果用傳統試題呈現,由於學習者無法直接觀察現 象,即使運用圖形來輔助測驗者理解題意,仍然有可能使得題意不清楚,影響 測驗的準確度。(二)將真實世界的情境於試題中表達出來,易使試題過於冗長。 譬如傳統試題測驗熱脹冷縮的原理時,要評量學習者是否能進行學習遷移,會 設計日常生活中的實際情況,讓學習者判斷並回答問題。題目爲了承載相關的

因此,近來有部分國際大型測驗,像是國際數學與科學潮流探究(TIMSS) 與國際學生評量計畫(PISA)都相繼採用科技結合試題呈現的方式,來改良傳 統測驗呈現試題上的缺失。譬如在測驗熱脹冷縮的原理時,PISA所呈現的試題 就直接給學生觀賞一段將熱開水淋在凹陷乒乓球的影片,片中呈現動態的膨脹 過程,讓真實生活的情境能夠展現並藉此測驗學生的相關概念。即使如此,部 分現象藉由影片仍然不易呈現,例如正斷層或逆斷層。因此,藉由動畫的設計 可以彌補此一缺點,並使得試題更容易爲測驗者所理解,從而能更準確的測量 學生的學習狀況。

談到動畫, Cook (2006)明確的指出, 動畫能夠在一段時間呈現多重圖像, 閱聽者透過觀賞動畫, 能看見一個現象完整的動態過程。有時候, 某些大尺度 的現象, 譬如板塊的移動, 或者微小尺度的現象, 如細胞的減數分裂, 因為不

易直接觀察,所以動畫就可以扮演相當良好的媒體,來呈現事件的過程。

因此,將動畫融入試題設計主要有下列三個優點:首先,在動畫試題中, 動畫扮演輔助的外部表徵,並能提供比圖像更豐富的訊息,讓測驗者能更具體 的了解題目所要敘述的抽象概念或現象,像是在地球科學領域中有關於漲潮與 退潮的現象。其次,Dancy & Beichner (2006)指出,動畫如同影片,相較於 靜態的圖片,動畫可以更有效清楚的呈現真實世界的情境,譬如颱風的移動路 線。最後,動畫如卡通般的外觀,能有效的提升學生的學習動機(Wouters, Paas & van Merrienboer, 2008)。

為了改善傳統試題設計的缺失,本研究以地球科學為領域,參考相關學者 的研究建議設計結合科技的動畫試題,旨在期望動畫試題除了能夠改良傳統試 題的缺失之外,還能因為融入相關的科技而提升學生測驗的意願與動機。放眼 當下,動畫與試題結合的相關研究仍然付之闕如,特別是在地球科學教育領域。 因此,也期望透過本研究的探究過程,將這個有效並有趣的試題設計作一初探, 彌補當前地球科學評量相關研究缺乏的情形(註1)。

貳、動畫設計理論基礎

動畫可以具體的呈現科學原理的運作或問題解決的方式,學習者透過其豐富的元件以及互動關係,得以理解複雜的科學原理或技能。然而即使動畫有助於輔助抽象的認知思考,動畫的設計必須配合學生的認知結構,並設法減少學習者的認知負荷。設計多媒體或動畫並用以傳遞知識,可能形成的認知負荷, 文獻指出,主要分爲三個種類,包括本質上的認知負荷(Intrinsic cognitive load)、外部的認知負荷(Extraneous cognitive load)以及適切的認知負荷 (Germane cognitive load)(Paas, Renkl, & Sweller, 2003)。

本質的認知負荷,起源於知識本身內部結構的交互作用,而越複雜的技能 或知識,因爲其內部知識結構相互關係更密切,越容易產生高度的本質認知負

荷。至於外部的認知負荷,則來自於訊息或知識的呈現過程;外部的認知負荷, 也可被稱爲無效的認知負荷(Ineffective cognitive load),設計不良的動畫或多 媒體,特別容易產生這類的認知負荷。很多時候,學習者面對動畫,往往在學 習知識之前,得先運用有限的認知資源去理解動畫呈現過程中的部分缺失。近 來常被研究的現象,也就是分散注意力效應(Split attention effect),與外部認 知負荷的關係密不可分。最後談到適切的認知負荷,也被稱爲有效的認知負荷 (Effective cognitive load),顧名思義,這種形式的認知負荷有助於學習,因爲 它能促進基模的建立並且能夠將基模自動化。

本研究的目的就是希望透過動畫來協助學生減少外部的認知負荷,在工作 記憶有限的前提之下,因減少外部認知負荷所釋放出的工作記憶將可用於長期 記憶當中基模的建構與自動化,也就是上一段提到的適切認知負荷。換句話說, 參與測驗的學生將能因動畫試題而減少因理解題目過程所產生的認知負荷,進 而更有效的運用認知資源來解決問題。

儘管設計得當的動畫能有效降低學習者認知負荷,但仍有學者指出,多數 的動畫因太過複雜或者呈現的速度太快,造成學習者無法正確的接收動畫所傳 遞的訊息,而無法配合學生的認知結構,將造成原本連續呈現的事件在學生眼 中轉成零碎、分離的片段,阻礙學生學習與理解(Tversky, Morrison & Betrancourt, 2002)。因此,若能依據下列相關研究所提出設計動畫的相關準則,就能避免動 畫對學習者造成認知負荷的超載。

一、互動性(Interactivity)

Cook(2006)與 Tversky, Morrison & Betrancourt(2002)都認為互動性(Interactivity)是設計優良的動畫所不可或缺的元素。透過互動,學生可以依據他們的認知需求與學習節奏來控制動畫的播放,讓他們能更正確的接受動畫所傳遞的資訊。

二、省略細節(Design with less detail)

一般而言,爲了減少學習者訊息處理的負荷,設計動畫時應該將原本複雜 的物件簡化,或者應避免多餘的裝飾或不相關的干擾,以凸顯主題,讓學習者 省去交互尋找比對(cross-searching and cross-comparing)的困擾。

三、設計輔助提示(Adding the annotation)

輔助提示的元件包括箭頭、反白、閃耀等引導的設計,如此將可以引導學 習者的注意力轉至動畫的關鍵處,減少因搜尋所產生的認知負荷(Tversky, Morrison & Betrancourt, 2002)。

總而言之,如果動畫能夠依據上述互動性、簡潔性以及引導性的原則來設計,將能有效的減少閱聽者的認知負荷,進而使其能運用更多的認知資源來理 解與學習。

參、動畫試題(Animation-based test)

依據上述的理論,本研究所使用的動畫皆考慮相關的設計準則進行設計, 並將這些動畫分別加入測驗的題目當中,成爲題目的一部分。試題的概念涵蓋 了地球科學領域中的重要觀念,包括斷層活動、洋流、潮汐、氣旋與反氣旋、 颱風、地球的自轉與公轉等概念。圖1是動畫試題的其中一道題目。題目中雙 黑線代表鐵路,山脈也以簡單的圖形呈現;此外,動畫有安排引導箭頭與播放 重播等互動元件。藉由這些設計,期望測驗者能因此操控更多被釋放出的認知 資源。

必須強調的是,強調互動性、簡潔性與引導性的動畫試題,依據理論,最 主要的優點,在於動畫能夠有效的減少學生答題時所產生的認知負荷。原因是 設計良好的動畫,在學生面對問題時,可以適時的提供外部的導引(External guidance)來彌補學生原來可能缺乏的相關基模;相對於動畫試題,傳統圖文 題呈現靜態的題目,學生必須自行進行運思,包括自行尋找或提取基模以做為內部導引(Internal guidance)來解決問題,因此會耗費較多的認知資源,造成較大的認知負荷。因此理論上,面對圖文題時,相較於動畫試題,學生會有較大的認知負荷。

當然,亦有文獻提到,高知識學生腦中存有比低知識學生更多的基模,因 此處理問題時高知識學生能藉由基模把複雜的運思程序簡化成小單位,也因此 高知識學生理論上不需要外部的導引,就可以順利且在認知負荷不超載的情況 下解決問題。在這種情況下,若再提供動畫給高知識學生,可能使他們要進行 額外的交叉比對,比對外部導引與既有的內部導引,此一認知活動理論上會增 加高知識學生的認知負荷。所以動畫試題可能的缺點之一,就是對於高知識的 學生產生額外的認知負荷。



圖1 動畫試題範例

在操作試題方面,當學生按下播放動畫時,題目就會出現一段動畫,譬如 有一股洋流緩緩的經臺灣海峽往北流的動畫,或者一個颱風正慢慢接近南台灣 的屏東等。學生作答時將答案填入線上測驗卷,共有二十道題目,當題目全部 回答完畢,受試者按下送出鍵,系統就會將該學生的答案傳送至研究者的線上

資料庫,有效克服時空的藩籬。

相較於動畫試題清楚呈現動態現象,傳統圖文題則以靜態方式表現題意, 並輔以箭頭,因此動畫試題與傳統圖文題無論在題目內容、題目順序、圖片素 材上都相同,唯一不同之處,就是動態與靜態的題目呈現方式。圖2爲圖文題 之範例。透過兩張圖可以發現,圖文題以箭頭呈現中洋脊斷層的錯動情形,相 較於動畫試題以動態呈現斷層的移動,靜態圖文題就如同傳統的考試卷,僅提 供靜態的圖片表徵。也因爲以靜態呈現,所以傳統圖文題沒有互動的元件,受 試者並不能操控題目。

本研究之動畫試題由五位地球科學專家與一位地球科學教育專家進行專家 效度審查,而動畫試題的信度係數為 0.7,難度則為 0.5。



3 此為中洋脊的斷層錯動情形,請判斷此轉型斷 層為下列何者? a. 上移運動 b. 下移運動 c. 左移運動 d. 右移運動



圖 2 圖文試題範例

肆、實驗結果

本研究主要目的為進行前測(pilot study)來初步探究受測者的反應並依據 實驗結果修正試題。研究對象選取臺北市某一高中高一學生 194 位,這些學生 均已修過一個學期的地球科學課程。在所有受測學生當中,有 114 位被隨機分 配到動畫測驗組,另外 80 位學生被分配到傳統圖文組。傳統圖文組的題目與動

畫組唯一的差別,就是呈現試題的方式,至於題目的排列順序、測驗內容、選 項安排等都沒有差別。施測時間進行約一個小時,而在測驗結束後,為了取得 相關質性資料以輔助量化資料,針對 19 位受試學生進行晤談,藉此了解他們對 於進行創新測驗方式的感受與態度。

值得一提的是,由於先備知識會影響學習者理解動畫等外部表徵,因此本 研究分別依據學生於上學期的地球科學段考成績來將學生分為高先備知識、中 先備知識與低先備知識。劃分的標準為高先備知識學生分數高於平均分數一個 標準差或以上,低先備知識學生則低於平均分數至少一個標準差,其餘學生則 分類在中先備知識。動畫組與圖文組的學生均依此先備知識標準分為低中高三 群。

研究結果顯示,依據理論基礎並考量認知負荷理論所設計的動畫試題,對 學生產生了部分的效益。根據質性資料分析,學生表示動畫題組比起傳統圖文 題組,讓試題變得更容易理解(Comprehensible)、更有趣味性(Interesting), 並感覺減少了認知試題所產生的負荷。在平均成績方面,動畫組的平均分數為 57.2分,圖文組為52.9分,動畫組的平均分數顯著高於圖文組的平均分數(*t*=1.91, *p*=.05, *d*=0.3, small effect size)。(表 1,圖 3)值得注意的是,當我們比較兩組 低先備知識的學生,動畫組的學生平均分數為53.9分,而圖文組則為43.5分, 動畫組的低先備知識學生表現優於圖文組低先備知識學生(*t*=1.94, *p*=.06, *d*=0.7, approaching large effect size)。

表1 動畫組與圖文組依先備知識之平均分數比較

先備知識	動畫組			圖文組					
	n	М	SD	n	М	SD	t	Р	d
總計	114	57.2	15.03	80	52.9	15.99	1.91	0.05*	0.3
高	30	62.3	13.87	13	60.9	16.95	0.25	0.80	0.1
¢	65	56.7	15.23	49	52.8	14.95	1.44	0.15	0.3
低	19	53.9	14.51	18	43.5	15.05	1.94	0.06	0.7

* p<0.05





圖 3 動畫組與圖文組依先備知識之平均分數盒狀圖

伍、討論

本研究結果顯示,面對敘述較為抽象的測驗問題,動畫能有效協助學生進 行運思,特別是針對先備知識較低的學生,這種輔助的效應更爲明顯。這樣的 發現與 Mayer (2001)的研究發現是一致的。Mayer (2001)指出,多媒體輔助 教材(包括動畫)的效益,對於低知識的參與者較爲明顯,相較之下,高知識 的參與者獲得動畫的幫助較爲有限,稱爲個別差異原則(Individual differences principle)。之所以會如此,主要是由於高知識的學習者單藉由圖文就足以形成 用以理解問題所需的心智表徵(mental representation),而低知識的學習者往往 在運用工作記憶形成心智表徵上有其困難,故須要動畫的外部表徵協助。而動 畫試題的功能,就是提供一個外部的鷹架(scaffold),協助受試者形成一個完 整的內部表徵,然後受試者就能參照這個表徵來解決問題,減少因題意不明所 產生的阻礙。

從認知心理學的角度來看,設計得當的動畫可以清楚且適切的表現動態過 程,此一特性有助於減少受試者的外部認知負荷(extraneous cognitive load), 讓受試者能更充分的運用其有限的認知資源。對於本次實驗中的低先備知識學 生而言,動畫無形之中減少了他們原本在理解題意時所需運用的工作記憶量, 讓他們有更充裕的資源用於問題解決,因此動畫組的低先備知識學生得分會高 於圖文組的低先備知識學生。對於中先備知識與高先備知識的學生而言,動畫 的效果就較不明顯,因爲先備知識高代表學生長期記憶中已存有先前建構好的 基模(schema),因此這些學生能隨時提取這些基模來解決問題,也因爲這些基 模的存在,無論是否有動畫,高先備知識的學生在運用工作記憶上皆能藉著使 用基模來減少工作記憶的負荷,造成無論在動畫組或圖文組的中高先備知識學 生,平均分數都相距不大。

陸、結論

本次研究的結果發現動畫試題對於學生在理解試題與答題態度上皆有提 升,而這樣的效果對於低先備知識的學生尤其明顯。而本研究的定位是初探性 質的研究,藉由這個實驗,我們會進一步依據答題的結果修正題目,提升其信 度與效度,並將太容易的題目修改爲稍具難度,以提升本試卷的鑑別度。此外, 將來的研究也會進一步增加樣本人數,並依據本研究的發現,將部分對學生造 成高認知負荷的題目置換,俾使試題能更符合學生的認知需求。最後,本研究 探討動畫的效能並結合在試題呈現上,初步結果符合預期,同時也說明了動畫 試題有其發展性,也期望未來的試題開發技術能更進一步考量受試者的認知需 求,如此將能有效提升測驗的準確性與預測性。

註 1:本文部分內容已發表在英國教育科技期刊(British Journal of Educational Technology),詳情請 查閱參考文獻。

參考文獻

- Cook, M. P. (2006). Visual representation in science education: the influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, 90, 1073-1091.
- Dancy, M. H. & Beichner, R. (2006). Impact of animation on assessment of conceptual understanding in physics. *Physics Education Research*, 2, 010104.
- Mayer, R. E. (2001). Multimedia Learning. Cambridge, UK: Cambridge University Press.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38, 1-4.
- Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Does it facilitate learning? *International Journal of Human-Computer Studies*, 57, 247–262.
- Wouters, P. & Paas, F. & van Merrienboer, J. J. G. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load, *Review of Educational Research*, 78(3), 645-675
- Wu, H. C., Yeh, T. K., & Chang, C. Y. (2010). The design of an animation-based test system in the area of Earth sciences. *British Journal of Educational Technology*, 41(3), E53-E57.colloquium

考試學刊 第九期 2011.06 Bulletin of Testing and Assessment

高中生英語聽力測驗之可行性探究

游春琪

大學入學考試中心

摘要

大學入學考試中心(簡稱大考中心)向教育部申請補助,探討具有重要決 定或高風險性的評量(high-stake test)之英語聽力測驗(簡稱英聽)實施之可 行性探究,本研究透過文獻探討了解亞洲地區大學入學考試英聽實施方式,同 時進行專家訪談與大型問卷等方式,以了解各大學校系對於英語文能力之需求 及增加英聽對於高中教學之可能影響。研究結果顯示,受訪者(大學部分 70% 與高中部分 65%)贊同英文考科增列聽力題型,並且認爲藉由英聽的實施將對 教學產生正面的影響。

關鍵詞:英語聽力測驗、高風險性評量



Investigating the Feasibility of Implementing an English Listening Comprehension Test for High School Students

Chun-Chi Yu

College Entrance Examination Center

Abstract

This report presents the findings of a research project commissioned by the Ministry of Education, which aimed to investigate the feasibility of implementing an English Listening Comprehension Test in high-stake tests in Taiwan. The research utilized document analyses, cross-country comparisons, expert consultation, questionnaire surveys, seminars, and special interest group discussions. The results of the survey indicated that approximately 65% of high school teachers and 70 % of universities supported the use of English listening comprehension test in high-stakes testing.

Keywords: English listening comprehension test, high-stakes testing

Chun-Chi Yu, Senior Staff Member, College Entrance Examination Center

壹、前言

2009年12月,Educational Testing Service(ETS)在台辦事處公布2008年 全球多益測驗(Test of English for International Communication, TOEIC)成績, 臺灣考生全體表現不如日韓,受到當時行政院副院長朱立倫的重視,甚至認為 是「奇恥大辱」。教育部長吳清基則強調,「台灣不能輸!」,將責成教育部 全盤檢討國小到高中的英語課綱,絕對要有效提升學生的英語能力,也會請大 考中心改進命題技術,讓老師的教學更有效(王彩鸝,2009)。

教育部歷來頒布之普通高級中學英文課程綱要,均以增進聽、說、讀、寫 正確英語文的能力,以應用於實際生活之溝通為首要目標。1973 至 1980 年 間,大學暨獨立學院入學考試(以下簡稱大考)英文考科試題均為選擇題,考 生作答方式仍限於閱讀試卷文字,尙未評量聽力與寫作能力,1981 年起英文 考科增列翻譯與英文作文,寫作能力開始納入大考的評量項目¹。

綜觀與台灣地緣、文化背景較爲相近之中國大陸、香港、日本、韓國等 地,其大考英語(文)考科之考試內容(表 1),均以閱讀能力爲主軸,且包 含聽力題型,而中國大陸及香港,更包括了聽、說、讀、寫四類能力的評量。

2010年1月,行政院舉行「全國人才培育會議」,會中為加強世界語言能力,研議大學指定科目考試(以下簡稱指考)英文考科增列聽力及口說題型; 同年2月財團法人大學入學考試中心基金會(以下簡稱大考中心)向教育部申 請補助「大學多元入學方案之檢討與改進研究計畫」,其中有三項子計畫與語 文能力的測驗有關,分別是:「本國語文寫作能力測驗(國文寫作能力測驗) 研究」、「英文寫作能力測驗研究」與「英文聽力測驗之可行性探究」。²「英

¹ 民國 100 年1月14日大學招生委員會聯合會(招聯會)第5屆常務委員會第7次會議決議: 贊成大考中心試辦英語聽力測驗,未來各招生管道可視其招生需求與學系特色,自行決定是 否參考英聽成績。大考中心將於100年9月試辦「大學入學考試中心高中英語聽力測驗」。

²計畫主持人為簡茂發教授,協同主持人為沈青嵩教授與洪冬桂教授,研究人員包括:李振清 教授、張武昌教授、陳浩然教授、蕭次融教授、林秀慧小姐、游春琪小姐。

文聽力測驗之可行性探究」計畫的主要研究目的為:

一、了解各大學校系對於英文能力(聽、說、讀、寫)測驗成績之需求性。

二、了解增加英語聽力測驗對於高中教學之可能影響。

三、了解其他國家英語聽力測驗之實施方式。

表1 亞洲各國(地)大學入學考試英語文考科評量之語言能力

語言能力 國別/地區	聽	說	讀明	寫
台灣	-	-	\checkmark	~
中國大陸	\checkmark	\checkmark	\checkmark	\checkmark
香港	\checkmark	\checkmark	\checkmark	\checkmark
日本	~	-	~	-
韓國	~	-	~	-
新加坡	大學入學考試未評量,但於 中學入學考試評量。		\checkmark	\checkmark

貳、文獻探討

本研究除了以華人社會的台灣、中國大陸、香港、新加坡為主要探究範圍 外,也包含地緣關係及歷史文化與華人較為相近之日本與韓國。

一、我國英語聽力測驗相關研究與實施

1980 年以前,大考的英文考科試題均為選擇題,雖稱評量考生聽力、閱 讀、文法、寫作等能力,考生作答方式仍限於閱讀試卷文字,尙無法確實評量 聽力與寫作能力。自 1981 年起英文考科增列翻譯與英文作文,配分比例各為 20%,寫作能力評量於焉開始。之後,教育界開始提出評量聽力能力的重要性 及相關研究。

(一) 政策推動

1997 年,國民中學課程標準即積極推動聽力培養,教育部委託國立教育 資料館³,依據國立編譯館編製的教科書內容製作聽力錄音帶,並提供各校使用。

教育部於1998年提出,八十八學年度大學推薦甄選第二階段,是否加考英語聽力,授權各校自行決定⁴。

國立臺灣師大附中於臺北區八十八學年度(1999)辦理高中推薦甄選入學, 於第二階段其他項目甄試之英語科試題中,已包含聽力題型(短句、對話、短 文),占整卷30%⁵。

2010年1月行政院舉辦之「全國人才培育會議」,建議大學指考之英語考 科增列聽力及口說題型,以加強世界語言能力。

(二)測驗研究建議

大考中心在「我國大學入學制度改革建議書—大學多元入學方案」(1992) 研究計畫中,建議指定科目考試(以下簡稱指考)英文考科加考聽力,以提供 相關院系參考使用。

黃燦遂等(1997)針對大學對於學生英語文能力的需求進行問卷調查,結 果顯示,大學各學院均不定期舉辦以英語發表的演講,約90%要求學生需具備 聽的能力。同時,在回收的問卷中,90%的英外文系,50%的文法商理工農醫 學院,希望於指考中加考英語聽力。

鄭恆雄等(1999)舉辦大學聯考與學科能力測驗(簡稱學測)英文科命題 探究座談會,多數與會學者建議加考聽力。

股允美等(2008)執行「95 課綱試題研發工作計畫—英文科(96)」,針對 95 課綱相關議題進行問卷調查,在 205 份有效問卷中,63.5%的高中教師認

³ 國立教育資料館及國立編譯館已於 2011 年 3 月 30 日整併入教育部之國家教育研究院。

⁴ 参見:http://www.edu.tw/content.aspx?site_content_sn=6329。

⁵ 參見:http://content.edu.tw/junior/english/sc_edu/testd/8803.doc。

為有必要增加其他語言能力,這些教師中有90.7%認為須增加聽力測驗。

大考中心(2009)召開研究發展委員會第 51 次會議暨命題研究委員會第 82 次會議聯席會議中,非英語相關科系教授建議加考聽力測驗,認為將有助 提升國人英語文能力。

大考中心於 2010 年舉辦「新修訂課綱(99 年) AB 分版,大學入學考試測 驗內容與命題方向之因應-英文科座談會」中,建議英語聽力測驗能列於相關 考試中,以期提升高中生英語聽力能力,並銜接未來大學之教育。

(三) 其他相關研究

謝國平、施玉惠、周見賢(1990)自1988年起接受教育廳委託,進行三年 的專題研究:「高中聯考英語科聽力測驗技術問題探討」,研究發現,若平日 教學經常舉行英語聽力測驗,考生較能適應聽力考試;學校裡之安靜教室與吵 鬧教室對於考生成績差異不大;但若平日未舉行過英語聽力測驗,在安靜教室 中作答的考生成績比吵鬧教室中的考生表現好。

鄭恆雄等(1992)提出編製英語聽力測驗之相關命題及試務原則。

蕭次融(1998)提出加考英語聽力的步驟:命題研究、訂定英語聽力成績 代表意義及使用方式、與大學校系溝通,促其招生採納英語聽力成績、向高中 師生說明英語聽力測驗、出版認識英語聽力類的參考說明。

黃燦遂等(2000)訂定聽力測驗命題參考原則,根據參考原則設計試題, 共含四種題型:單句、對話、短篇報告、新聞稿及演講稿,其中單句及對話皆 以圖畫輔助,試題並進行研究測試,通過率皆達 50%。

二、亞洲地區大學入學考試的英語聽力實施方式

以下以 2010 年為例,說明亞洲主要地區,包括台灣、中國大陸、香港、 日本、韓國以及新加坡等地之大學入學考試英語(文)考科中,英語聽力題型 的實施方式,詳見表 2,各詳細內容請參閱附錄一。

國別/地區	台灣	中國大陸	香港	日本	韓國	新加坡
考試	學測 指考	高考	高級會考	大学入試 センター 試験	大學修學 能力試驗	GCE A-Level
目的	升學	升學	升學 就業	升學	升學	升學
日程	1/29~30 7/1~3	6/7~8	3~6月	1/17~18	11/13	5~11 月
應考 人數	142,129 83,927	9,494,747	38,964	512,451	670,000	
測驗實施 方式		廣播或 CD音響	廣播	IC Player	廣播	廣播或 CD音響
首次實 施年度		2001年	1986年*	2006年	1993年	1984 年**

表2 亞洲主要地區2010年大學入學考試英語(文)考科聽力測驗實施方式

註* 於中學會考首測。

** 於 GCE N-Level 首測。

(一)中國大陸

中國大陸舉行的大學入學考試稱為「普通高等學校招生全國統一考試」(以下簡稱高考),圖1顯示中國大陸在高考發展中推動英語聽力測驗的重要歷程。首先,1993年推行「全日制普通高級中學英語教學大綱」,測試部分包含筆試和口試或聽力,藉以帶動中學加強英語師資及改善語音設施。1993至1996年教育部考試中心在湖北省進行之高中會考中增加聽力部分,證實會考中加設聽力的可行性和必要性。1997至1999年教育部考試中心開始在廣東省進行高考中增加聽力的可行性研究⁶,先於1997年廣東高考英語試卷增加聽

⁶ 1996 年絕大多數省市的中等學校高中階段招生考試之英語試題均加入聽力,配分至多 30 分至 少 15 分,平均占全卷的 18%左右。(http://www.ncneedu.cn/old/Teaching/View.asp?id=5362)。

力,但不計入總分,僅做參考,其後於 1998 年設計聽力測驗占整卷 10%,至 1999 年聽力測驗更占整卷的 20%。大陸教育部考試中心並於 1999 年高考英語 試卷中增加聽力部分,且制定過渡方案。2000 年各省提供含聽力及不含聽力 的兩類試卷。2003 年各地開始統一使用含聽力部分的高考英語試卷。2005 年 外語聽力不再做統一要求,各省自行決定外語聽力測試的考試形式、時間和計 分。有些省分,如浙江和雲南,將全國英語等級考試 (Public English Test System,簡稱 PETS)第二級中的聽力成績納入高考英語科總分,以減輕高考 的試務負擔,也提供考生多次參加考試的機會(劉慶思,2010)。在 2010 年, 大部分省、市、自治區的高考試卷皆評量聽力,但新疆、四川、河北、陝西的 試卷不含英聽(新疆於 2011 年恢復聽力考試⁷),山西、河南、吉林考聽力, 但不計入總分,天津則於 3 月考,可考兩次,取最高分。



圖1 中國大陸英語聽力測試推動歷程

目前,高考之外語學科共提供英、日、俄、德、法、西六個語種供考生選擇,每個語種均提供含聽力和不含聽力兩類試卷供各省分選擇。以省(自治區、直轄市)為單位,由各省的考試中心以中國教育部公布之普通高等學校招

⁷ 參見: http://big5.china.com.cn/gate/big5/edu.china.com.cn/2011-02/13/content_21910642.htm。

生全國統一考試大綱為核心,除以教育部考試中心所命試題外,部分省分亦自 主命題,允許在考試科目、試題與計分方式有所不同,全國共二十份試卷。

測驗方式上,聽力測驗乃透過學校廣播系統播放試題,每間教室皆備有完善的視聽設備與監試系統。各省、市、自治區考生人數中,屬河南省最多 952,400人,最少的為西藏 18,028人,詳細考生人數請見附錄一。

中國大陸對於英語聽力考試的重視程度,可從報紙新聞觀之,其標題「孩 子考聽力 全社會在幫忙 地上沒喇叭 天上沒雷響」,如同全民運動,考場外考 生家長自發成立「護考督察隊」,密切監控考點周邊「聲音環境」,確保孩子 能在最安靜的環境裡接收到最清晰的聽力信號。集體關閉手機,減少頻率干擾 8。

(二)香港

1986年,香港中學會考(Hong Kong Certificate of Education, HKCEE)英 文科首次引入聆聽考試(即英聽考試)⁹,多年來,香港考試及評核局(簡稱考 評局)仍持續提升測驗時播音與收音的品質。最早第一代聆聽考試採用無線感 應迴線系統,考生使用由考評局提供的無線耳機收聽試場內的錄音廣播。此方 法需要大量耳機,常因器材耗損,令耳機接收出現問題;另外,在貯存、維修 保護和收發器材方面支出龐大¹⁰。

1995 年聆聽考試開始以大氣電波播放聆聽資料,改由考生自備收音機及 耳機到場應考。當時由負責英文節目的香港電台第四台直播,但因經常被投訴 接收不清,後改由訊號較強的香港電台第二台直播。在播放聆聽錄音前,電台 會播放倫敦小交響樂團的《綠袖子幻想曲》,讓考生放鬆心情。每個試場會安 排一間「特別室」,考生如接收不清,可以到採用中央廣播的特別室繼續收聽 及考試。但有考生埋怨,他們在前往特別室的途中會錯過部分聆聽資料,影響

⁸ 參見:http://sh.xtata.com/s/20100609/48127328.shtml 。

⁹ 參見: http://www.hkeaa.edu.hk/doclibrary/about_hkeaa/annual_report/ar2007_section1.pdf。

¹⁰ 参見:http://www.com.cuhk.edu.hk/ubeat_past/100597/hkcee.html。

作答。

部分考場因地理位置而導致電台訊號接收不佳,考評局於 2010 年引入不 易受干擾的紅外線接收系統。語音訊號由光碟播放機傳遞至安裝在禮堂的紅外 線發射器(圖 2),再透過紅外線訊號傳遞到試場內考生的無線耳機,考生無 需使用收音機接收信號。不過紅外線適宜近距離傳送,只能在無障礙的地方使 用,不能穿透牆壁¹¹。



圖 2 香港英聽考試使用紅外線發射器

(三)日本

日本文部科學省於 2003 年制定《培養「能使用英語的日本人」行動計畫》 (The Action Plan to Cultivate "Japanese with English Abilities"),明訂 2006 年 大學入學考試(大学入試センター試験)加考英語聽力,爲該計畫重要目標之 一。

¹¹ 參見: http://orientaldaily.on.cc/cnt/news/20100328/00176_005.htmlkcee.html 。

英語科考試屬外語考科之一¹²,由考生選考,其中報考英語者自 2006 年起 須另加考一堂聽力測驗。考試時,每位考生提供一台 IC Player,如圖 3。考試 時間 60 分鐘,前 30 分鐘爲考試說明(含機器使用說明),後 30 分鐘爲作答時 間,當年於原報名費中增加日幣 2000 円。



圖 3 日本大學入試 IC Player

(四)韓國

韓國於 1993 年大學修學能力試驗(College Scholastic Ability Test, CSAT) 加考英語聽力,採廣播方式播放試題,測驗時間為下午 1:05 ~ 1:35。為了 完成聽力測驗,也幾乎全員動員,約 125 班次班機更改起飛時刻,所有火車禁 鳴喇叭(緊急時刻例外)。考試當天地鐵與公車皆增加班次,並且於考場周圍 200 公尺內禁止停車,以防止交通擁塞;上班時間延後一小時,由原上午 9:

¹² 外語考科含英、德、法、中、韓語。

00 調整為 10:00。相關的金融交易時間也隨之更改¹³。

目前,韓國預計於 2012 年推出 The National English Ability Test(簡稱 NEAT),評量聽、說、讀、寫四種語言能力。測驗方式仿托福形式,屬網路 測驗(Internet-based Test, IBT),暫分三等級:第一級,以成人為對象,測驗 職場用英語能力,96題,作答限時 137 分鐘(相較於托福 104 題,225 分鐘; 多益 200 題,120 分鐘。),由多家大企業協助編製。第二、三級,皆以高中 以上學生爲對象,前者測驗升大學、出國留學、或外國學生來南韓以英語修讀 學位所需英語能力;後者測驗日常生活英語會話能力。第二、三級作爲大學入 學用、第一級爲大學二或三年級畢業考試或就業用,規劃將於 2016 年取代 CSAT 中的英文考科,考生於平日參加 NEAT,其成績可以作爲大學入學使用 ¹⁴。

(五) 新加坡

新加坡之英語聽力測驗首次於 1984 年 The Singapore-Cambridge General Certificate of Education Normal Level (簡稱 GCE N Level)施測,當時因在保密環境下,無法確保錄音帶的製作品質,各考場以廣播方式經由收音機播放試題(Tan, Chow & Goh, 2008)。

新加坡教育學制為:小學教育(六年)、中等教育(四年或五年)、後中 等教育(二或三年)及大學(三或四年)。小學畢業時(約12歲)參加 Primary School Leaving Examination(簡稱 PSLE);中等教育畢業時(約15歲),須 參加 GCE N (Technical) Level(簡稱 GCE N (T)-Level)、N (Academic) Level (簡稱 GCE N (A)-Level)或 Ordinary Level(簡稱 GCE O-Level);後中等教 育畢業時(18-19歲),須參加 GCE Advanced Level(簡稱 GCE A-Level),這

¹³ 参見: http://www.koreatimes.co.kr/www/news/nation/2008/11/117_34287.html; http://www.koreatimes.co.kr/www/news/include/print.asp?newsIdx=76611。

¹⁴ 参見: http://www.kice.re.kr/ko/download.do?attach_id=18684; http://www.koreatimes.co.kr/www/news/nation/2011/01/286_80560.html;與 http://www.koreatimes.co.kr/www/news/nation/2010/09/113_72541.html。

些考試中之英文考科,皆評量聽、說、讀、寫語言技能(圖4),但聽力主要於 PSLE與 GCE N(T)-Level 評量,使用廣播或 CD 音響。GCE A-Level 考試僅 評量閱讀與寫作,不含聽力;然而,在評量口說能力時,其實已同時評量了聽力。



圖 4 新加坡各級英語考試中評量各種能力之比例

綜觀上述各主要地區進行大學入學考試英語聽力測驗的施測經驗,若台灣 也要針對大學入學舉辦英語聽力測驗,應注意以下問題:

1.考生人數與英語聽力施測方式

大學入學考試攸關考生未來的升學與就業,因此首重公平性,也使得任何 要進行大規模施測的計畫常令人卻步。台灣英語教學界自 1989 年起即展開英 語聽力相關研究,至今已達二十多年,然而卻遲遲未能進行正式施測。綜觀亞 洲各國,新加坡早於 1984 年開始施測,最晚的日本也已於 2006 年開始施測, 而考生人數,則由最少的香港三萬多名至中國大陸的近千萬名之多,至於施測 方式,採用校園廣播、個人化 IC Player 或是以收音機播放。

2.大規模考試引發的爭議

中國大陸從要求各省實施多年後,開放各省自行決定是否實施英語聽力測驗,進而促使各校提升視聽設備,直接影響日常教學;香港 2010 年增設紅外線語音接收系統,以期解決各種接收不清問題,然而,也有考生反映試題播放品質不佳;日本重視個人化,使用 IC Player,試題僅能播放一次,考後考生可自行帶回,然而,自 2010 年起,均由日本入試中心於考後統一回收;韓國考試期間,為提升播放品質,影響人民日常作息,多少須付出社會成本。

綜觀亞洲各國實施英語聽力測驗的經驗,仍有相當多問題需要解決,例 如:聽力設備故障、考場周邊瞬間電磁波干擾,廣播發生幾秒鐘的停頓、供電 故障、錄音品質問題、天氣影響如雷擊、考場外圍環境影響,如汽機車、火 車、飛機、人員吵鬧聲等,還有試務工作人員失誤、考生及家長聯名要求加分 或重考聽力等問題,皆需考量完備的配套措施與有效的因應對策。因此,台灣 在實施英語聽力測驗時,各國的經驗將極具參考價值。

參、研究方法

為深入探討針對臺灣高中生實施英語聽力測驗之可行性,大考中心 2010 年再組成專案研究小組,以能確實落實測驗為前提,在 10 個月研究期間,首 先蒐集文獻,了解我國曾經進行過之英語聽力相關研究,以及亞洲其他國家英 語聽力測驗之實施方式;其次,以小型座談會與諮詢訪談方式完成問卷設計以 進行問卷調查,同時針對英語聽力及英文寫作辦理焦點座談會與大型研討會 議,以便蒐集與諮詢各相關領域之專家與學者意見。研究流程如圖 5 所示。

為廣泛蒐集大學端與高中端對高中生實施英語聽力測驗的意見,本研究以 網路問卷方式,調查大學與高中對相關議題的看法,問卷填答時程爲民國 99 年6月29日至7月16日。問卷內容包括大學校系對大學生英語聽力之能力要 求,包括聽、說、讀、寫;高中針對英語聽力的教學情況;還有,若大學入學

研究方法 2月 與設計 3~5月 文獻探討 專家訪談 座談會召 問卷設計 6~8月 與官施 開 8、9月 間卷分析 專家諮詢 10、11月 報告撰寫

考試加考英語聽力測驗對高中教學的影響等,詳細問卷內容如附錄二。 以下簡述本研究問卷結果之分析,及專家學者意見之彙整。

圖 5 研究流程

肆、研究結果分析

一、問卷調查分析

爲更深入瞭解大學選才端與高中教學端對聽力測驗的看法,分別設計針對 大學教師與高中教師問卷(問卷結果請參閱附錄二),大學問卷內容涉及招生 及大一新生英語文能力相關議題;高中問卷包括高中教學現況、學測與指考英 (語)文評量功能、英語聽力等相關議題。採用網路問卷方式,對象以大學教 授或熟悉招生事物之同仁及高中教師為主,以了解英語聽力測驗之可行性。

大學部分共發函 1468 個科系,邀請各系系主任或具代表意見者,彙整系上 意見後,再行上網填寫問卷。高中部分共寄發 382 所高中,邀請各校至少一位 英文科教師填寫問卷(填寫時須反映該校綜合意見)。大學方面的問卷回收率 58%(約 851 份);高中方面的問卷回收率 54%(約 205 份)。



問卷填答分析以量化分析為主,並在各項量化分析後羅列質化的個別討論 或說明。

(一)大學問卷

各大學校系對於大一新生應具備的英語文能力顯示,以閱讀比例最高 (92%)、其次是聽力(75%)、接著是口說(56%)、然後是寫作(48%)。

而大一新生應具備英語文能力之原因中,閱讀英文教科書或相關參考資料 爲最多,占91%,接著搜尋網際網路占57%,聽懂英文演講占46%;另外,與 外籍同學互動、適應全英語上課內容、撰寫英文報告等皆各占40%以上。

30%認為該系的新生已具備該系所要求英語文能力,66%則認為新生尚未 具備該系所要求的英語文能力。

於各系的甄選入學中,除了參考學測成績外,32%仍於第二階段舉辦其他 英語文能力之相關測驗:英文閱讀、英語聽力、英語會話、英文寫作等項目。 目前國內外雖已有很多評量聽、說、讀、寫的英語文檢定測驗,然而對於大考 英文考科,有74%認為應再增加寫作、聽力、口說能力題型,其中77%認為需 增加聽力、56%認為需增加口說、51%認為需加重英文寫作比例。另外,若於 學測中提供考生分項能力成績,79%表示將會採計或參酌。

2010年1月行政院舉辦之「全國人才培育會議」中,建議大學指考之英文 考科增列聽力與口說題型,70%表示贊同,27%不贊同。國內外雖已有很多評 量聽、說、讀、寫的英語文檢定測驗,仍有 74%認為指考應增加英語聽力 (77%)、口說能力(56%)的評量,同時應增加英文寫作比例(51%)。

(二)高中問卷

如課程綱要所述,高中教師認為高中畢業生的英文能力皆須具備聽、說、 讀、寫的技能(各項比例皆超過95%);另外,也提及應具備國際觀、批判性 思考(critical thinking)能力、整合能力、及欣賞文化與文學的能力。

至於高中平日教學於聽、說、讀、寫的排列順序隨著年級而有所調整(圖 6),閱讀能力於高中三年皆受到極大的重視,聽力與口說的重視程度隨著年級

而減弱,寫作的重視隨著年級增加而加重,可能因爲受到學測與指考未評量聽 力與口說能力的影響。



圖 6 高中三年教學重視排列順序

雖然學測與指考未包含聽力題型,然而,高中校內考核項目若加入「聽力」 或「聽與說」,75%認為能促使聽、說、讀、寫並重。

進一步了解教學對於聽的訓練,有83%於教學中已涵蓋了聽的訓練,16% 未涵蓋聽的訓練,其原因有:58%認為因為教學內容繁複、52%因學生程度落 差大、36%因學生人數太多、18%認為會加重學生課業負擔,其它15%則是授 課時間不足及大學入學考試未考聽力,故教學都以閱讀為主。

在假設僅少數大學科系需要聽的能力時,77%高中教師仍會加強學生的英語聽力。

聽力題型於日常教學中的施測,94%表示曾經實施,其中最常於小考、期 中考、期末考中實施。



英文教學中,多媒體設備(如電腦、MP3等)的使用頻率高達 92%、傳統 音響設備(如 CD 音響)也有 60%。

在實施聽力題型測驗時,常造成困擾的因素為:因機器或系統播放不良與 教室隔音不佳各占 45%與 44%、學校缺乏完善的錄音設備,影響試題錄製品質 占 23%、學生座位前後影響與常有飛機或汽車等噪音干擾,各占 18%、12%。

2010年1月行政院舉辦之「全國人才培育會議」中,建議大學指考之英文 考科增列聽力與口說題型,65%表示贊同,33%不贊同;若加考英語聽力題型 時,88%顯示會改變教學、11%不會改變教學。不會改變教學的原因,除了教 學時間不足與學生人數太多等因素外,很多的教師提及該校或個人平日教學早 已重視聽力。

當問及學測或指考英文考科應包含的英語文能力(聽、說、讀、寫)評量, 以閱讀能力比例最高97%、其次是寫作能力占96%、接著爲聽力70%、口說占 37%。

二、諮詢與相關會議意見彙整

為了解各界對英聽實施可行性的想法與意見,本計畫舉辦二次大型座談 會、多次專家諮詢、多次研究小組會議,討論議題主要為:

1. 各大學校系對於英語文測驗成績需求性

2. 增加英語聽力測驗對於高中教學之可能影響

將其意見彙整如下,提供未來正式推出聽力測驗時之參考:

(一)大學教學現況

全國大學幾乎所有的領域皆已推動英語授課,而非僅於外文學系。加入英 語聽力考試,高中生才會重視,大學則會因爲高中已經加重了英語聽力的訓 練,大學在推動英語授課的政策時,較無阻力,學生會更進入狀況。

(二) 高中教學現況

讀和聽稱為被動/接受性技能(receptive skills),說和寫稱為主動性/能產

性技能(productive skills),語言的學習過程應為:聽→說→讀→寫。加考聽 力,應可更實際培養考生聽、說能力,並提升寫的能力;學生語言能力進步, 進而減少教師教學的負擔,朝向教學優質化邁進。藉由聽力測驗與教學,啓發 學生學習潛能;因此學生可提升語言能力的發展,有助改善學生的學習策略。 另外,1979年大學聯招英文考科提出加考英文作文,社會反彈聲浪大,但漸 漸地對高中教學產生正向的回沖效應(washback effect)。

英聽實施,學生應較不排斥,因一般學生的聽力較寫作及口說能力理想。 目前國中三年中,段考聽力部分比例各約為 30%、20%、10%,逐年減少之 因,乃為國民中學基本學力測驗未含聽力題型。高中階段,高一和高二之學校 期中、期末考、小考早已實施,但到了高三就不再評量聽力,主要因為大考不 考英語聽力,老師和學生認為沒有繼續的必要。雖然考試引導教學並非最佳學 習策略,但是不容置疑,大考加考英聽,學生學習的動機將改變。加考英聽應 不至於造成高中教學端極大反彈,但需要時間的適應與緩衝。

英聽推出前,建議提供長遠準備期,以利教學因應,另外,也需增加教師 的研習機會。

(三) 高中教學設備

現行高中教科書皆隨書附光碟,除了上課外,學生也可利用時間自己練 習。對於偏遠地區是否有視聽設備短缺的問題,應實地了解,請相關當局補足 資源。據觀察,絕大多數學校視聽設備都相當足夠,也很先進。考試機構要務 是將考試辦好,擬定相關配套措施並落實,至於教學端的問題,必須由教學端 設法解決;學習端,則是確實掌握自己的學習部分,各司其職。

(四) 城鄉差異

城鄉差距在每個地區都有,比如中國大陸,甚至大台北地區也存在些許的 差距,因此不應歸咎在城鄉差距。英聽實施後才能真正解決城鄉差距的問題, 不應以此爲藉口,改變考試模式,才能改變教學的模式,各界才能重視。


(五) 英聽實施建議

試題的設計必須採情境式、生活化,以學生學習的範疇為試題的內容,才 能達到學習與考試相互配合。試題必須是直接的(direct),而非僅測驗慣用語 (idioms)或單詞聽寫。

試題的難度要能兼顧不同能力群考生,試題難易度比例需適度分布,難易 皆有,才能有效鑑別考生的能力。

外界多建議大考若須包含英聽,可使用全民英語能力分級檢定測驗,但需 思考其範圍是否符合「普通高級中學課程綱要」、是否能因應課程綱要修訂而 隨之調整測驗內容、及二者測驗目的是否相同等。

英聽的比例可先占整卷10%,之後再逐漸增加比例。待實施順利,再考慮 增加較高階的聽力評量。英聽可採選考方式進行,才不至於造成太大的反彈聲 浪,另外,並非所有考生都需要英聽的成績。

英聽應採獨立施測,增加考試次數,考生自行選考,但一年多試的缺點是 學生要多次準備;然而,就分散考試壓力而言,多次較有彈性,壓力較小,較 符合需求。選才端,大學各校系可自訂門檻,英聽成績的使用可以更多元。

英聽在實施方式上採取漸進式,待外界肯定後再進行全國性統一考試。大 考中心若推出英聽,大學校系將無須另外檢定學生語言程度。

伍、結論與建議

一、結論

亞洲各國英語聽力測驗已實行多年,各國均認為提升學生英語文能力,加 考英語聽力測驗有其迫切性,台灣若要推出聽力測驗,各國多年實施經驗將可 供參考。

根據本研究的問卷調查結果,大學學測或指考英文考科增列聽力題型:在 大學部分,有 70%贊同,27%不贊同;而高中部分,有 65%表示贊同,33%不

贊同;若加考英語聽力題型時,88.29%高中教師認為會改變高中教學、11%則 表示不會改變教學,因平日教學中已包含聽力。

二、建議

高中教師在平日教學中,英語聽力測驗的評量已實施多年。加考英語聽力,將更促使教師與學生對英語聽力的重視,進而提升考生英語文能力。因爲 學習管道並非僅有閱讀,聽力也是吸取知識之重要途徑,但需要時間的適應和 緩衝。

正視城鄉差距問題,實施英語聽力,才能真正了解問題癥結,進而解決問題,城鄉差距不能一直被當作藉口。改變考試模式,才能改變教學模式;換言之,大學入學考試中,增加英語聽力測驗,各界才會對英語聽力加以重視。

試題部分,應明訂聽力測驗範圍、題型,並配合高中課程綱要,試題設計 需考量學習者能力差異。英語聽力試題的難易度控制也須確實掌控。

在實施做法上,可採取漸進式方式,分時實施及一年多試,測驗成績可作 爲選才參考,而非立即列爲學測或指考中單獨考科成績計算。先以試辦方式進 行,待一切理論與實務問題都可解決,同時本方案也獲致外界肯定,再行研究 是否納入全國統一考試的一部分。



參考文獻

- 大學入學考試中心(1992)。我國大學入學制度改革建議書-大學多元入學方案。台北: 作者。
- 大學入學考試中心(2009)。研究發展委員會第51次會議暨命題研究委員會第82次會 議聯席會議紀錄。台北:作者。
- 王彩鸝(2009,12月18日)。學生英文差朱立倫:奇恥大辱吳清基:台灣不能輸! 將檢討課綱、考試項目。聯合晩報,A4。
- 謝國平、施玉惠、周見賢(1990)。高中聯考英語科聽力測驗技術問題之研究(三)。南 投:省政府教育廳。
- 股允美、楊泰雄、葉錫南、林秀慧、游春琪(2008)。95 課網試題研發工作計畫-英文 科(96)研究報告。台北:大學入學考試中心。
- 黄燦遂、周中天、陳秀英、陳昭芬、陳坤田(2000)。**指定科目考試規劃研究期末報告** (英文科 II)。台北:大學入學考試中心。
- 黃燦遂、陳秀英、林雪娥、楊貞婉、陳坤田(1997)。**指定科目考試規劃研究報告(英 文科)**。台北:大學入學考試中心。
- 鄭恆雄、湯廷池、黃自來、吳耀敦、廖樹輝、斯定國、陳坤田(1992)。**大學入學考試** 英文試題編製、預試及聽力研究報告。台北:大學入學考試中心。
- 鄭恆雄、楊懿麗、陳須姬、陳凌霞、陳坤田、吳惠珠(1999)。**大學聯考與學科能力測 驗英文科命題探究座談會報告**。台北:大學入學考試中心。
- 蕭次融(1998)。大學入學考試加考英語聽力。選才通訊,42。
- 劉慶思(2010)。中國大陸高考外語學科中的聽力測試。2010「海峽兩岸教育考試評價 改革發展研討會」。
- Tan, Y. K., Chow, H. K., & Goh, C. (2008). Examinations in Singapore: Change and continuity, 1891-2007. Singapore: World Scientific.
- 中國教育考試網 http://www.neea.edu.cn/
- 日本大學入試中心(National Center for University Entrance Examinations)
 - http://www.dnc.ac.jp/
- 香港考試及評核局 http://www.hkeaa.edu.hk/tc/
- 韓國教育課程評價院(KICE, Korea Institute for Curriculum and Evaluation) http://www.kice.re.kr

Singapore Examinations and Assessment Board http://www.seab.gov.sg/index.html

附錄一

亞洲地區大學入學考試英文考科說明,皆以2010年為例。

一、台灣

- •考試:學科能力測驗
- •辦理單位:大學入學考試中心
- •考試目的:大學入學用
- •考試日程:1/29~1/30
- •考生人數:142129人
- •作答時間:100分鐘
- ●評量語言能力:讀(68~72%)、寫(28~32%)
- ●題型:選擇題(四選一)與非選擇題
- •計分:15級分
- •考試:指定科目考試
- •辦理單位:大學入學考試中心
- •考試目的:大學入學用
- •考試日程:7/1~7/3
- •考生人數:83,927人
- •作答時間:80分鐘
- ●評量語言能力:讀(68~72%)、寫(28~32%)
- ●題型:選擇題(四選一)與非選擇題
- •計分:100分

二、中國大陸

- •考試:普通高等學校招生全國統一考試
- 辦理單位:考試中心
- •考試目的:大學入學用
- •考試日程:6/7~6/8
- •考生人數:9,494,747人



•作答時間:120分鐘

(聽:20分鐘、英語知識運用與閱讀:60分鐘、寫作:40分鐘)
•評量語言能力:聽(20%,30分)、讀(57%,85分)、寫(23%,35分)、
說(口說評量僅限於報考英語文相關科系,考試時間各省
另行安排。)

- •題型:選擇題(四選一)與非選擇題
- •計分:150分



圖 7 各省分考生人數

三、香港

- ●考試:香港高級程度會考(Hong Kong Advanced Level Examination)
- 辦理單位:香港考試及評核局
- •考試目的:大學入學及就業
- •考試日程:3月底至5月初
- •考生人數:38,964人

- •作答時間:350分鐘
- •評量語言能力:英(語) 文科目共二科:英語運用及英語文學,其中英語 運用評量聽(18%,60分)、寫(18%,75分)、讀(18%, 90分)、說(18%,20分)、運用(28%,105分)。
- •題型:選擇題、非選擇題
- •計分:6級

四、日本

- 考試:大学入試センター試験(The National Center Test)
- •辦理單位: The National Center for University Entrance Examinations
- •考試目的:大學入學用
- •考試日程:1/17~1/18
- ●考生人數:512,451人
- ●作答時間:80分鐘(讀)、60分鐘(聽)
- •評量語言能力:讀聽
- ●題型:選擇題(四選一)
- ●計分:200分(讀)、50分(聽)

五、韓國

- •考試:대학수학능력시험(大學修學能力試驗, The College Scholastic Ability Test)
- •辦理單位:Korea Institute for Curriculum and Evaluation
- •考試目的:大學入學用
- ●考試日程:11/18
- •考生人數:712,000人
- ●作答時間:70分鐘(聽:20分鐘、讀:50分鐘)
- ●評量語言能力:聽、讀、(寫)
- •題型:選擇題(五選一)
- •計分:100分



- 六、新加坡
 - •考試: The Singapore-Cambridge General Certificate of Education (Advanced Level) Examination
 - •辦理單位: Singapore Examinations and Assessment Board
 - •考試目的:大學入學用
 - •考試日程:5月底至12月初
 - •考生人數:14,280人(全體考生)
 - ●考試科目:General Paper H1
 - •作答時間:180分鐘(寫作:90分鐘、閱讀:90分鐘)
 - •評量語言能力:寫作(Paper 1)與閱讀(Paper 2)
 - •題型:非選擇題
 - •配分:寫作(50%)與閱讀(50%)



附錄二

一、大學問卷

 下列係大一新生應具備英語文能力的幾點主要原因。請勾選贊同的 選項?(可複選) 	比例
(A) 適應全英語上課內容	40.31%
(B)閱讀英文教科書或相關參考資料	90.60%
(C)觀看英語發音之專業影片	38.31%
(D)與外籍同學互動	40.07%
(E)撰寫英文報告	39.60%
(F)聽懂英文演講	46.06%
(G)參加國際研討會	35.49%
(H)參加國際交換課程	32.08%
(I)搜尋網際網路(Internet)參考資訊	56.76%
(J)其他	0.47%
未填寫	1.41%

 以貴系招生而言,大一新生應具備下列哪些面向的英語文能力? (可複選) 	比例
(A)聽力	74.74%
(B)口說	56.17%
(C)閱讀	92.01%
(D)寫作	47.94%
(E)其他	1.41%
未填寫	1.41%

 承上題,貴系大部分新生,是否已經具備貴系所要求的英語文能力? 	比例
(A)是	29.96%
(B)否	66.39%
未填寫	3.64%

- 115 📕

 在貴系的甄選入學,除了參考學測成績外,是否仍於第二階段舉辦 其他英語文能力之相關測驗? 	比例
(A)是,測驗項目(可複選)	32.08%
(1)英文寫作能力測驗	32.97%
(2)英語聽力測驗	41.76%
(3)英文閱讀測驗	62.27%
(4)英語會話測驗	41.03%
(5)其他	13.92%
(B)否	65.92%
未填寫	2.00%

5. 未來大考中心學測英文考科若提供考生寫作、聽力、口說分項能力	
之成績,貴系是否會採計或參酌,以作為甄選入學錄取學生的參	比例
考?	
(A)是	78.97%
(1)寫作	76.19%
(2)聽力	79.91%
(3)口說	67.26%
(B)否	18.68%
未填寫	2.35%

6.2010年1月行政院舉辦「全國人才培育會議」,會中建議大學指考	
之英語考科增列聽力及口說題型,您是否贊同?	
(A)贊同	69.45%
(B)不贊同	26.79%
未填寫	3.76%

考試學刊 第九期 2011.06 Bulletin of Testing and Assessment

 7. 目前國內外已有很多評量聽說讀寫的英語文檢定測驗,您認為指考 中應增加那些語言能力題型?(可複選) 	比例
(A)需增加	74.03%
(1)英文寫作比例加重	51.27%
(2)英語聽力	77.14%
(3)口說	56.35%
(4)其他	0.95%
(B)不需增加	23.74%
未填寫	2.23%

二、高中部分

1. 您認為高中畢業生需要具備下列哪些英語文能力?(可複選)	比例
(A)聽	99.02%
(B)說	96.59%
(C)讀	99.02%
(D)寫	98.54%
(E)其他	7.80%

2. ^z	平日上課聽	、說、讀、	寫之教學	,您的重礼	見排列順序:	爲何?	
	年級	聽	說	讀	寫		
	一	30.85%	19.35%	32.93%	16.87%		
	吉一 同—	25.32%	16.69%	35.89%	22.10%		
	高三	21.00%	13.55%	35.75%	29.70%		

3. 您在平日的英文教學中時,使用的教學設備與使用頻率為何?	
(A)傳統音響設備(如 CD 音響)	60%
(B)多媒體設備(如電腦、MP3等)	92%
(C)其他	21%
未填寫	19%

- 117 🔳

4.	雖然目前學測與指考僅評量閱讀與寫作能力,您在平日的教學中是	
	否涵蓋聽的訓練?	
	(A)是	83.41%
	(B)否,原因:	16.10%
	(1)會加重學生課業負擔	18.18%
	(2)學生人數太多	36.36%
	(3)教學內容繁複	57.58%
	(4)學生程度落差大	51.52%
	(5)其他	15.15%
未	填寫	0.49%

5.	若學測或指考增加聽與口說的題型,您是否會改變目前的教學方	
	式?	
	(A)會	88.29%
	(B)不會	10.73%
未	填寫	0.98%

6. 您認爲學測或指考應包含哪些英語文能力之評量?(可複選)	
(A)閱讀能力	97.07%
(B)寫作能力	96.10%
(C)聽力	70.24%
(D)口說能力	37.07%
(E)其他	0.98%
未填寫	0.49%

7.2010年1月行政院舉辦之「全國人才培育會議」建議大學指考之英	
語考科增列聽力及口說題型,您是否贊同?	
(A)贊同	64.88%
(B)不贊同	33.17%
未填寫	1.95%

8.	如果只有少數大學科系需要聽的能力,您會加強訓練學生的英文聽	
	力嗎?	
	(A)會	76.59%
	(B)不會	22.93%
	未填寫	0.49%

9.	學測與指考未含聽力題型,在校成績考核項目中,若加入「聽力」	
	或「聽與說」之項目,您認為將會促使聽、說、讀、寫並重嗎?	
	(A)會	74.63%
	(B)不會	24.88%
	未填寫	0.49%

10. 貴校實施聽力評量情形如何?	
(A)從未實施	6.34%
(B)曾經實施	93.66%
(1)小考	36.79%
(2)期中考	32.51%
(3)期末考	29.12%
(4)模擬考	1.58%

11. 貴校進行聽力題型測驗時,最困擾的因素為何?(可複選)	
(A)因機器或系統不良播放,導致情況不佳。	44.88%
(B)教室隔音不佳各班播放試題時會相互影響	43.90%
(C)因地緣關係常有飛機、汽車等噪音干擾	11.71%
(D)學校缺乏完善的錄音設備,試題錄製品質(雜音、音量控制)	22.93%
不佳	
(E)學生座位前後影響聽的效果	17.56%
(F)其他	16.10%
未填寫	7.32%

考試學刊 第九期 2011.06 ■ 120 Bulletin of Testing and Assessment

歷史試題命製的基本觀念

張 元

清華大學

摘要

課堂教學有其目標,能否達成目標需要檢測,考試是檢測的主要手段。歷 史課堂教學的目標,應是帶領學生初步理解歷史知識的性質、結構與方法,而 不是記憶過去發生的事件。

「歷史」一詞,原有兩層意思,一指過去發生的事件,一指對過去事件的 理解與表述。課堂教學的重點是後者,而非前者,這是教師應該銘記於心的重 要觀念。我們時常提及的「歷史思維」,簡單來說,就是對過去的理解;培養學 生的思維能力,可以從引導學生如何理解過去的事件著手。

按照此一教學目標設計的考試,命題的重點應在測試學生的歷史思維能力,也就是檢測學生是否具有這門課程提供的思考能力。本文以 98、99 年大學入學指定考試歷史考科某些考察考生思考能力的試題爲例,逐題作一簡述,並特別著力於對「閱讀」式試題的分析。最後,筆者提出設計此類思考型試題的具體建議,也命製若干例題,以供讀者參考。

關鍵詞:歷史科評量、歷史教學、試題分析、歷史思維能力、歷史指考



張 元,清華大學歷史研究所榮譽退休教授

Basic Concepts of History Test Item Writing

Yuan Chang

National Tsing Hua University

Abstract

There are objectives of classroom teaching. Examinations serve as a primary means of testing whether students achieve these objectives. The goal of history teaching should lead students to a preliminary understanding of the nature, the structure, and the methodology of historical knowledge, rather than the memory of past events.

Originally, "history" has two meanings: one refers to the events that happened in the past; the other means understanding and interpretation of past events. Teachers should bear in mind that it is the latter, not the former, that relies on class teaching. For instance, the term "historical thinking" refers to the understanding of the past. To cultivate students' skills in "historical thinking," teachers can start with assisting their students to understand past events.

In accordance with the teaching objectives, the test design should focus on the student's ability of historical thinking. That is, a test should be designed to examine whether the students have acquired the ability of historical thinking. In this paper, the DRT history test items of 2009 and 2010 academic years are adopted and the test items related to "historical thinking" are specified. These test items are briefly interpreted. The main focus is on the test items based on "reading." Finally, the author proposes some specific suggestions for writing the test items based on "reading," and sample test items are also provided.

Keywords: history assessment, history teaching, item analysis, historical thinking/thinking skills, History Subject in DRT



Yuan Chang, National Tsing Hua University, Honorary Retired Professor

前言

所謂「基本觀念」,指一些最基礎的認知,關係到所做事情的每一步,所以, 十分要緊。不過,這種要緊的「觀念」,往往欠缺共識。彼此之間,也存在著明 顯的差異。需要經過一番討論,或許可以出現為大多數人接受的觀點。我所提 出的,僅是個人的想法,也請大家指教。

壹、教學與考試的關係

「考試領導教學」是大家常聽到的話,考試或許正在領導教學,但考試應 該領導教學嗎?稍稍想想,就會覺得不可以。為什麼?教學有其目標,教學活 動應該朝向這個目標,怎麼可以把「考試」當成目標呢?那麼,課程的教學目 標又是怎樣擬定呢?我覺得,任何課程的教學目標都應是把學生領進這門課程 的知識領域,讓學生有初步的接觸與認識。那麼,這門課程的知識領域又指什 麼呢?我想,不應該是知識的具體內容而已,知識的性質、結構和方法,可能 更為重要。特別是歷史,學生應該知道的,與其說是過去發生的事情,不如說 是我們怎麼知道過去的這些事情。這是一個重要的觀念,我將在下一個小節討 論。

老師朝著依知識領域擬定的教學目標進行教學,成效如何?需要檢測,這時就用到考試了。所以,我們常說考試是查驗教學是否取得成效的方法或手段。 這樣說來,我們平時對學生進行的各種「考試」,應該以我們教課的內容為目標, 看看學生的理解程度,而不是模擬諸如「學測」、「指考」試題加以練習。也就 是說,學測、指考的試題不應該扮演領導我們教學的角色。

老師會說,我們要學生多練習,目的只希望他們在重要的考試中取得好成 績。這個成績,對學生來說非常重要,家長、校長也都很看重,我們的壓力不 小,輕忽不得。這是實情,學生在重要考試的成績,影響很大,不容小覷,作

為學科老師,必須負責。然而,我們是否想到另外的方法,可以兼顧二者;既 可以達成教學目標,又有助於學生取得重要考試的好成績。在理論上,這是可 以做到的。因為理論上學生的學習應該在於增強他們的學科「能力」,而不是對 於瑣碎內容的記憶。就歷史科而言,學科能力無非指對過去有一定程度的理解、 知道如何表述的能力,而不是對過去事件的記誦能力。所以,在理論上,只要 我們在課堂上不斷加強學生對歷史知識的理解,並予以表述的練習,他們的學 科能力必然有所增強,就足以應付重要考試,並取得優異成績。

事實如何?是另一問題。如果,理論上應該如此,事實上難以執行;那麼, 可能的關鍵,應該是老師對於「歷史知識」的認知仍然不夠清晰,基本觀念未 能有所掌握。

貳、歷史一詞的兩層意思

歷史知識的最基本觀念是「歷史」這個詞彙,有兩層意思,我們一定要清 楚掌握。何兆武在《歷史與歷史學》(香港:牛津大學出版社,1995)一書的「自 序」中第一句話就是:「通常我們使用的「歷史」一詞包含有兩層意思,一是指 過去發生過的事件,一是指我們對過去事件的理解和敍述。」英文的歷史教學 書籍,往往也把同樣的意思放在第一頁的第一句話。這句話非常重要,我們可 以作一個簡單的分析。歷史是過去發生的事,無人反對,但過去的事我們怎麼 知道呢?這就要靠歷史家的工作了。歷史家搜集資料,加以論證,作出敍述, 才是我們讀到的「歷史」,也就是歷史家對過去事件的理解與敍述。所以,我們 可以說,沒有歷史家,就沒有歷史;我們又可以說,認識歷史之前,要先認識 歷史家。何兆武在這篇自序的最後一句話寫道:「並不是有了活生生的歷史,就 會有活生生的歷史學;而是有了活生生的歷史學,然後才會有活生生的歷史。」 這就是說,我們的歷史知識來自歷史家所從事的歷史學工作,而不是來自從那 些過去曾經發生過的事件。

老師們若對歷史一詞的兩層意思未能深刻掌握,會有怎樣的情況?首先, 以爲課程內容全是「過去發生的事件」,老師在課堂上講的就是過去發生的事 件,如:秦始皇統一天下、焚書坑儒,漢武帝獨尊儒術、討伐匈奴等,中學的 歷史課這樣講,看來好像沒什麼不對;但是,這樣的講課,就會讓老師和學生 同樣感到歷史就是這些,如同教科書中所寫,記誦起來,就學會了。其次,如 果老師忘了歷史一詞的第二層意思,教課時就一直作史事的陳述,或增補一些 教科書中沒有的史事。整個課程就失去了的思辨性,就是在課堂上對學生提出 問題,也都屬於記憶性質,如什麼人?哪一年?等。歷史教學中很重要的「歷 史思維」,難以展現。第三,老師在閱讀與備課的時候,只會注意到史事的積累, 而未能注意到知識取得的過程。也就是說,老師在看書的時候,會注意那些新 的知識內容,可以在課堂上講述的新材料,而未能注意這本書的作者如何運用 資料、分析論證等屬於歷史學的地方。於是,命製試題時,未曾從這方面著手, 看到這一類試題,也難以帶領學生進行思考。我總覺得,只是因爲輕忽了一個 很重要的觀念,下了許多工夫,卻未見成效,實在是很可惜的事。

老師在教課的時候,會想到這段史事是從哪裡來的?是哪一本書,或哪位 史家嗎?如果想起來,一定要講出來,這是重要信息。讓學生知道,我們的歷 史知識是有來源的,是經由歷史家的工作,讓我們知道了這些。如果,我們是 從一本不夠嚴謹的歷史通俗讀物看到的,我們必須聲明在先;如果我們是從歷 史小說或歷史劇知道的,那最好是不要用於課堂,因爲那與歷史學無關。我們 都知道卡爾說過:「在研究事實之前,先該研究歷史家。」他說的是「研究」, 那麼我們在課堂教學,需要介紹歷史家嗎?我的想法是,我們不妨選擇最重要 的歷史家,作最簡要的介紹;至於一般的歷史學者,我們沒有時間介紹,也沒 有介紹的必要。

當我們說到這個講法來自那一位大史家時,其實已暗示了那位大史家為什 麼要提出這個說法,其中就有了「問題」的出現。我們就在講述一個問題的提 出與解答,而不是在陳述一個「事件」,這樣我們就把學生帶進了一個思考的過

程。我們講述一件事時,說到我們根據的資料來自何處,非但說明了資料與事件的關係,也會把這段資料的「意義」講了出來。我們知道,事件的本身是沒有「意義」的,任何意義都是後人加上去的;而我們教學的重點,不是沒有意義的事件本身,而是後人加於事件上的意義。由於我們講述的重點是歷史事件的意義,我們同樣可以帶領學生進行思考的練習。我們常說,歷史是一門思考的學問,不是記憶的學問;我們只有在課堂教學中帶領學生進行思考,才能把這句話加以落實。這也就是爲什麼我們要介紹大歷史家的緣由了,唯有大歷史家,方能把資料或事件的意義作最好的表述,選材精要,論證嚴謹,意境高遠, 文字優美,值得我們欣賞,更值得我們學習。我心目中的現代大歷史家,有: 王國維、陳寅恪、陳垣、錢穆,以及張蔭麟、湯用彤、郭廷以、范文瀾等。當然,這是十分主觀的意見,僅供參考。

參、思考與記憶

我們都知道,大歷史家的工作,主要展現於對資料的深入掌握,以及論證 的嚴謹。這些我們在課堂教學時,或可偶一介紹,但檢測學生是否理解,甚為 困難;也就是說,若在考題中呈現,要考出學生關於論證的能力,似乎稍早了 一點,目前並不恰當。不過,我們仍然應該把檢測學生對於歷史問題的論證能 力,作為未來考試目標中的重要項目。

歷史家在做論證之前,必須具備優異的閱讀能力,這是歷史學工作的基礎, 也應該是我們可以讓學生經由練習而取得的能力。不論是原始史料,或現代學 者的論著,所謂的二手史料,「閱讀」均非易事,都要經過練習,方能有所理解。 「閱讀」困難之處在於它往往包含了一個思考的過程,要通過思考、辨析方能 瞭解文字中的意思,才能說讀懂了,有了正確的理解。

相對於閱讀,就是記憶。也許我們可以很粗淺地把前者歸於歷史一詞的第二層意思,而後者就是第一層意思。我們要學生記得很多很多,某一方面就是

學生不必動用他們的思考能力;學生自然而然就覺得歷史就是一門記得過去事 件的學科,是不需要用腦筋的。所以,歷史是十分無聊的科目。

我們如何檢測學生的「思考」能力呢?「閱讀」式的試題就是一種方法。 閱讀式試題的特點在於不需記得什麼,但要有分析、思辨的能力;這是與「記 憶」式的試題必須記得什麼方能解答,有著基本上的差異。

近兩年的大學指定科目考試試題¹中出現較多的「閱讀」性質的試題,我們 不妨舉出,並配合答對率及鑑別指數表²,略作說明。

【98年指考】

第19題:

19.英國詩人但尼生在1847年寫下一首詩:「男人在戶外,女人在 家裡/男人使劍,女人拿針/男人有腦,女人有心/男人發號 施令,女人順從遵行/若不如此,混亂必行」這首詩反映出十 九世紀中葉的哪一種觀點:

(A)下層階級因經濟改善,主張女性不必外出工作,可以賦閒 在家

P=全體到考考生答對率 Ph=高分組(前 33%)考生答對率 Pl=低分組(後 33%)考生答對率 Pa~Pe=依成績高低分為各佔 20%的五種能力組考生答對率 D=全體到考考生鑑別度 Ph-Pl T=多選題全對率 D1~D4=五種能力組考生鑑別度(D1=Pa-Pb...D4=Pd-Pe) 題號前有星號者為多選題,以得分率代替答對率

(多選題)得分率 = $\left(\sum \frac{該題答對答案數 - is 題答錯答案數}{is 題答案數 \times 考生人數}\right) \times 100$



¹98年及99年指定科目考試歷史考科試題,取自大學入學考試中心 http://www.ceec.edu.tw。

² 98 年及 99 年指定科目考試歷史考科試題之答對率及鑑別指數表,取自大學入學考試中心 http://www.ceec.edu.tw。各值說明如下:

- (B)平民階層因工業革命後工作機會大減,主張女性應該在家工作
- (C)中產階級主張女性不應該外出工作,認為男、女內外必須 有別
- (D)婦女階層因性別意識高漲,主張女性不需負起家庭經濟的 重擔

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
19	86	89	84	91	87	86	85	83		5	4	1	1	2

考生只需閱讀,理解文義,即可選出正確答案。如果讀到題幹中的「若不 如此,混亂必行」,就想到應該是選項中的「必須有別」,亦可作對。其實,我 不覺得這是不妥的設計,考生若能讀到這一點,也反映其閱讀的細心與專注。

第30題:

- 30.1934年《食貨半月刊》創刊,其〈編輯的話〉指出:「史學雖 不是史料的單純排列,史學卻離不開史料。理論雖不是史料的 單純排列可以產生,理論並不是原形一擺,就算成功了的。中 國社會史的理論鬥爭,總算熱鬧過了。但是如不經一番史料的 搜求,特殊問題的提出和解決,那進一步的理論爭鬥,斷斷是 不能出現的。」這位編輯要表達的是:
 - (A)史學研究即是史料學
 (B)史料學可以取代理論

 (C)理論無益於史學研究
 (D)理論建構離不開史料

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
30	92	99	82	99	98	97	94	74		17	1	1	3	20

考生如果讀懂了意思,就可以作出正確選擇。儘管如此,這題考的卻是一個關於歷史學中理論與史料之間關係的討論,相當重要。如果考生具有一定的 思考能力,這是一道容易的題目。

第23題:

23.兩位清朝官員談到台灣的械鬥。熊一本說:「兩類肇端,每在連塍(土溝)爭水、強割佔耕、毫釐口角,致成大衅。」福康安說:「因閩庄、粤庄彼此交錯,田業毗連,遂有構衅相爭之事。」就兩人的說法來看,分類械鬥的起因是:
(A)利益競爭 (B)習慣不同 (C)官員放縱 (D)語言隔閡

題號	Р	Ph	Pl	Pa	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
23	93	99	83	100	99	97	93	77		16	1	2	4	16

考生只要靜下心來,把題目仔細讀一遍,就可以看出正確選項與錯誤選項 中的距離是蠻大的,這是通過率最高的一題。

第24題:

- 24.新加坡獨立時,李光耀公開聲明:「我們不是華僑,我們只是 華裔,這一點無人可以更改。中國是中國,我們是我們。」此 句話主要的用意為何?
 - (A)中國的施壓引起新加坡人極度的不滿,想撇清與中國的關係
 - (B)新加坡長期受到英國的統治,因英國的施壓不得不做此聲明
 - (C)新加坡華人經過英國長時期統治,已喪失對中國文化的認同
 - (D)適應新加坡的民族與政治情勢,去除印度人與馬來人的疑慮

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
24	51	77	26	82	66	52	37	20		51	16	14	15	17

選項中的中國施壓、英國施壓,以及對中國文化喪失認同,這些意思,均 未見於題幹的資料文字中。所以,只要細心閱讀,稍加思考,略知新加坡當地

情況,應能答對。

第29題:

29.「政府管制外匯、黃金,以平抑物價。1942年,發行一億美金公債,每一美元定價法幣20元,不久黑市高至法幣450元,但 有關人士,仍可按官價購進,大獲暴利。1943至1945年之間, 政府抛售黃金,1944年至1945年,舉辦黃金存款,利益多落入 大戶之手。」這段敘述最可能和下列哪一議題有關?
(A)抗戰後期政府經濟政策與實施情形
(B)抗戰期間美元與法幣的匯率變化
(C)抗戰後期政府對銀行、金店、大戶的控制
(D)抗戰期間政府放任黃金的買賣

題號	Р	Ph	Pl	Pa	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
29	68	88	45	92	82	72	58	37		43	10	10	14	21

題幹文字所述,是以外匯、美元、黃金和法幣等情況,說明當時的整體情形;只談匯率變化和黃金買賣是不對的,也未涉及對銀行等的控制。所以,只要細心閱讀,瞭解文義,即可答對。

第36題:

36.甲書說:「宗教改革的主因是羅馬教會本身的腐化。」乙書說: 「宗教改革起於馬丁路德在1517年公開批評天主教會的腐化。」以下哪個敘述較為合理:

(A)甲著眼於遠因,乙著眼於近因

- (B)甲著眼於導火線因素,乙著眼於背景因素
- (C)甲著眼於個人因素,乙著眼於制度性因素
- (D)甲著眼於表層因素,乙著眼於深層因素

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
36	62	88	30	92	80	67	48	21		58	12	13	19	27

題幹兩段文字,分別關於「背景」和「原因」,這也是我們閱讀時必須能夠 清楚分辨的。選項的差距很大,考生只要用心,應可答對。

第34題:

34.下列四段有關中共對知識分子思想鬥爭的敘述,請按合於邏輯 的次序加以排列。(甲)當時毛澤東正想找理由掀起批判胡適 的運動,這個事件正好提供引爆點,毛澤東說俞平伯販賣胡適 思想;(乙)俞平伯是胡適考證《紅樓夢》的嫡傳弟子,認為 《紅樓夢》是曹雪芹自傳,並從此一觀點考證《紅樓夢》;(丙) 霎時之間,全國報章雜誌都是批判胡適的文章,由文藝而擴及 哲學、歷史等範圍,歷時長達一年之久;(丁)引起兩位不知 名的山東作家的批評,說他的考證方法,是「資產階級的唯心 論」,目的在隱晦《紅樓夢》批判封建主義的性格。 (A)甲乙丙丁 (B)甲丙乙丁 (C)乙甲丙丁 (D)乙丁甲丙

題	號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
3	4	67	90	39	93	84	72	57	30		51	9	12	15	27

題幹的敍述文字就要考生找出四段文字的「邏輯次序」,等於純粹測試考生 閱讀文章時的思考能力。四段文字中,有一個發展的過程,要找出第一句話, 此後的每一句話都有其轉接的詞語,如第二句始於「引起」,接著是「當時」, 最後一句應該有總結的意味。

第18題:

18.班固說:「古代天子建國,諸侯立家,從卿大夫到庶人,各有 等差。人們遵奉並服事上級,而在下位的人也不會有踰越的念 頭。」接著又說:「五霸,三王之罪人也;六國,五霸之罪人



也;四豪(信陵君等四公子),又六國之罪人也。」從班固這 兩段敘述我們可以看到怎樣的歷史概念?

- (A)因果概念,夏、商、周與春秋的結束,是由於在下位的人 沒有踰越的念頭
- (B)時序概念,從夏商周到春秋到六國抗秦再到戰國,時間次 序不可變更
- (C)證據概念,三王、五霸、六國、四豪,都是古代人們嚴守 身分等差的事證
- (D)變遷概念,從夏、商、周到戰國後期,呈現逐漸遠離古代 理想的趨勢

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
18	73	93	45	95	89	80	65	35		48	6	9	15	30

題幹部分作了白話文的改寫,語意應該是明白的。選項問題的一些名詞, 如因果、時序、證據和變遷等概念,考生可能有點陌生。不過,只要讀了題幹 中班固所說的話,應該可以感覺到那種對古代理想已經越來越遠了。選項中的 因果概念在班固的文句中不是沒有,但選項的解釋明顯錯誤,可以不選。

第25題:

25.下列兩段文字,見於何書?(甲)忽報孔明遣人送書至,周瑜 覽畢,長歎一聲,昏絕,徐徐又醒,仰天長嘆曰:「既生瑜, 何生亮?」連叫數聲而亡。(乙)周瑜還江陵,為行裝,而道 於巴陵病逝,時年三十六,權素服舉哀,感動左右。
(A)甲、乙皆見於《三國演義》
(B)甲見於《三國演義》,乙見於《三國志》
(C)甲、乙皆見於《三國志》
(D)甲見於《三國志》,乙見於《三國演義》

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
25	76	93	53	96	89	81	68	45		40	7	8	13	23

兩段資料,一段取自《三國演義》,一段取自《三國志》。只要學生能夠分辨通俗小說與正史的表述不同,其中有很大的差異,就能做出正確選擇。

第12題:

12.「宇文泰重用蘇綽,制定計賬、戶籍等制度;獎勵清廉,嚴禁 貪污,建議裁減官員,進行屯田,又作〈六條詔書〉,宇文泰 完全採納,令百官誦習。州刺史郡守縣官不通六條及計賬法, 不許做官。西魏政治顯然比東魏政治好一些,____。」 在空格處,應該接著的是哪句話最為適宜?
(A)宇文泰選取身體強健的農民組成府兵
(B)西魏後來成為北周,東魏後來成為北齊
(C)到孝文帝時進行大規模的華化運動
(D)宇文氏的政權從此趨於鞏固

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
12	46	62	32	67	53	45	38	28		30	14	8	7	10

題幹中一段較長的敍述,看看能夠得出怎樣的結語。我們可以把敍述當成 事例,看看可以支持怎樣的論斷;也可以把它看作有了這些事,或做了這些事, 將會得到怎樣的結果;還可以用套入的方法試試,看看哪一句讀起來最通順。 這是需要運用思考的,但要找到思考的路。考生只想史實會是怎樣,就會選到 史實正確,卻與問題無關的選項。學生不會做,大概是從未見過這樣的題型, 不知從何入手。

多重選擇四題之中,有三題屬於「閱讀」的試題。

第38題:



38.馮道是五代名臣,作〈長樂老自敘〉,敘述歷事四朝及契丹所得的階勳官爵以為榮,自謂為臣為子為父為師,都無愧色。後人對他有些意見,如甲:「他的道德如同古代聖賢,才能很高, 氣量很大,雖然朝代改變,大家對他均無批評。」乙:「我讀 馮道的〈自敘〉,看他所說的光榮,真可以說是無廉恥的人, 天下國家也就可想而知了。」丙:「馮道的行為,很有古人風 範;馮道的氣量,深得大臣的體要。」在下列敘述中,選出適 當者。

(A)三種敘述,評價相同,都是讚揚

(B)甲是批評,丙是事實描述

- (C)乙是讚揚,丙是嚴厲批評
- (D)甲是稱讚,乙是嚴厲批評

(E)觀點不同,評價有異,馮道即是一例

題號	Р	Ph	Pl	Pa	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
*38	90	98	75	99	97	95	90	67	87	23	2	2	5	23

題幹中有三段文字,是關於五代馮道的評價,每一段學生要看出是褒或是 貶,就能夠選出正確的選項。

第39題:

39.一位羅馬作家對時局有這樣的評論:「帝國之內一片昇平,官員的名號一如往昔。年輕一代曾目睹共和昔日面目者少之又少!」「國家已經脫胎換骨,昔日優美的道德已一絲不存。平等遭到剝奪,所有人都一股腦兒,仰望聽命於第一公民一人。」 我們如何解讀這份資料?

(A)這位作家遭逢從「共和」到「帝制」的轉變
(B)這位作家口中的「第一公民」指的就是凱撒
(C)「第一公民」與「奧古斯都」字義並無不同
(D)「昔日優美道德」是指羅馬的傳統公民精神
(E)從作家的言論推測,他應是一位共和主義者

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
*39	50	61	36	63	57	52	47	29	21	25	6	5	5	18

題幹中的兩段文字,說到羅馬從共和到帝國的變遷,書寫者顯然對於眼見 的變化很不以爲然,他還是緬懷昔日的時光。考生如能讀懂,也不難選出正確 的選項。這道題目分數稍低,仍有記憶成份,應爲原因之一。

第40題:

40.三位史家解釋西元七、八世紀阿拉伯帝國的擴張。史家甲:「阿拉伯人朝向單一目標:征服世界,使所有人皈依伊斯蘭教。」 史家乙:「穆斯林社區的統一可藉由對外攻擊而獲得維持,這 也會提升哈里發的權威。」史家丙:「一開始的大征服並非伊 斯蘭的擴張,而是阿拉伯民族的擴張。」以下何者為是:
(A)甲不同意宗教性因素
(B)乙強調是政治性因素
(C)丙不同意甲的解釋
(D)丙著眼於宗教性因素
(E)甲和乙的說法沒有差別

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
*40	73	89	54	91	84	77	68	46	60	35	7	7	9	22

題幹三段文字,主要呈現兩種概念,一是政治的,另一是宗教的。考生若 能清楚分辨,亦不難作出正確選擇。

這份試卷中,考生得分最低的一題,並不是「思考」類型的題目,而是「記 億」類型的題目。這是**第 21 題**:

21. 俄國人喜歡用不同的地名來反映政治情勢的改變,下列是同一城市的不同地名,請依時序選出正確的排列。
 (A)聖彼得堡—彼得格勒—列寧格勒—聖彼得堡



(B) 聖彼得堡—列寧格勒—史達林格勒—聖彼德堡

(C)彼得格勒—列寧格勒—史達林格勒—彼得格勒

(D)彼得格勒—聖彼得堡—列寧格勒—彼得格勒

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
21	10	16	7	19	10	8	7	7		9	9	2	1	0

這道題目的解答,考生一定要記得兩個「事實」,一是今天的聖彼得堡就是 與彼得大帝有關的那座城市,曾改名為列寧格勒,再改回原名。二是,昔日的 蘇聯,除了列寧格勒,另有一座著名的大城,叫史達林格勒。如果考生不知道 這兩件事,他就無法作出正確的選擇。從這一題與以上題目的比較,我們可以 很清楚看到「思考」型題目與「記憶」型題目的明顯差異。

這份試卷的選擇題部分,最大的特點就是「思考」型題目的大量增加。考 試的次日,各大報報導了這次指考的試題,值得注意的是,「老師」覺得歷史題 目很難,是「史上第二難」,考生卻反映不難、容易。何以如此?或許題型上的 變化,讓老師覺得不適應,而考生卻覺得省掉很多記憶,只要細心閱讀題幹資 料,加以思考,選出正確答案,並非難事。

99年的指考試題,這類「閱讀」型的題目,較98年為少,未能將此一新的題型繼續有所發展。

第6題:

6. 日本某歷史教科書,用「侵略」一詞形容二十世紀的中日戰爭。日本 文部省的審查意見認為:「侵略」是含有負面道德暗示的字眼,用於 下一代國民教育的教科書,來描寫自己國家的行動,是不妥的,應該 改用「軍事前進」之類的文字。對於日本官方這種做法,我們應如何 理解較為恰當?

(A)對於發動戰爭一事,日本政府並未能確切深入檢討與反省(B)對於發動戰爭一事,日本政府已從教育革新做起,知所省思

(C)歷史教科書的文字,應該屬於事實敘述,避免道德評價(D)日本政府藉此做法,企圖得到中、韓政府的諒解與同意

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
6	45	53	37	57	47	45	42	34		16	10	2	3	8

題幹文字十分清楚, 敍述日本政府對二十世紀的中日戰爭, 不願以「侵略」 一詞, 用於教科書的表述, 可視為未作確切的檢討與反省。日本政府的態度與 德國相比, 很是不同。題幹的敍述, 與教科書的文字, 究竟應屬事實或難免道 德評價, 並無關聯, 這是另一個複雜的問題。

第19題:

19. 英國史家霍布斯邦(Eric J. Hobsbawm)說:「若無十九世紀資產階級 社會的解體在先,勢無十月革命以及蘇聯的成立在後。」又說:「發 生於二十世紀30年代的大蕭條,使人們認為蘇聯的經濟制度成為一條 全球性的可行之路。」依據這樣的陳述,下列何者是最恰當的理解?
(A)資本主義得以挑戰社會主義,最大的力量來自於本身的強大
(B)社會主義得以挑戰資本主義,最大的力量來自於對手的弱點
(D)社會主義得以挑戰資本主義,最大的力量來自於對手的弱點

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
19	80	99	49	99	97	91	75	36		50	2	6	16	39

這是一道典型的思考題,題幹文字十分清楚,只要讀懂,略加思考,即可 答對。唯一需要具備的知識是「資產階級社會」即指選項中的「資本主義」;「蘇 聯」即指「社會主義」。



第25題:

- 25. 戊戌維新之前,主張維新的人,發表各自意見。嚴復在〈上皇帝萬言書〉請皇帝「結百姓之心」方法是「親至沿海各地,巡守省方,縱民嵩呼,瞻識共主。」譚嗣同在《仁學》中說:「君也者,為民辦事者也。事不辨而易其人,亦天下之通義也。」依據這些資料,下列敘述何者最為適宜?
 - (A)兩位維新者,都抱有新思想,都反對舊觀念
 - (B)認同維新者,重點各有不同,但都主張民主
 - (C)嚴復思想激進,譚嗣同思想保守

(D)嚴復思想保守,譚嗣同思想激進

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
25	43	70	16	76	58	42	25	12		54	18	16	17	13

題幹中的兩段文字中,考生讀懂嚴復是擁護君主制,屬保守思想;譚嗣同 則不贊同君主制,屬激進思想,即可答對。得分偏低,尤其是低分組,值得探 討。

第26題:

- 26.《史記》有〈河渠書〉,《漢書》有〈溝洫志〉,主要針對黃河患害的嚴重而寫的。但是,《後漢書》到新、舊《唐書》都沒有〈河渠書〉 或〈溝洫志〉等篇章。學者歸因於黃河基本上安流無事,無需專闢一篇。請問此一時期黃河安流的最可能原因是什麼?
 - (A)北方朝廷政治清明,治理黄河成效顯著
 - (B)北方胡人政權重用漢人,不乏水利專家
 - (C)黄河中游西側農業大為進步,水利發達
 - (D)黄河中游西側多屬胡人居處,變農為牧

題號	Р	Ph	Pl	Pa	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
26	29	46	15	51	35	27	19	13		31	16	8	8	6

題幹主要說明有一段時期「黃河安流」的史料依據,問的是此一現象的最 好解釋。學生需要利用今天生態學的觀念來思考,理解農業開發可能會破壞生 態,而放牧業則能保護生態;這一時期黃河安流,主因在於華北生態未受破壞, 變農爲牧之故。這是一道頗有現實意味的題目,考生一時未能領悟,也是常情。

第27題:

27. 一部編年史如此描述1980年代世界局勢的變動:「經過20年社會、政治的劇變,一種新的政治浪潮在西方世界逐漸抬頭。這可說是一種反革命,開始於1979 年柴契爾(Margaret Thatcher)夫人出任英國首相,但直到1980年底雷根(Ronald Reagan)當選美國總統後,這一浪潮才真正發揮力量。」文中所說「新的政治浪潮」所指為何?其主要政策路線為何?

(A)新自由主義;新中間路線(C)新保守主義;自由市場

(B)新社會主義;福利國家(D)新民主主義;人權政治

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
27	51	78	23	83	67	52	36	17		55	16	15	16	19

題幹所摘錄文字中,如果注意到「反革命」字眼,又對英國柴契爾夫人與 美國雷根兩位執政者的作為有一點印象,選出正確答案,應不困難。

第29題:

- 29. 三國時期,孫權要呂蒙多讀書,呂蒙讀了一段時間後,魯肅對呂蒙關於時局的分析,大為欽佩,並說呂蒙「非復吳下阿蒙」。根據以上所述,請問孫權要呂蒙讀什麼書?
 - (A)《詩》、《書》、《易》、《禮》、《春秋》,這些基礎典籍, 屬於學者必讀
 - (B)《論語》、《孟子》、《大學》、《中庸》等,明白做人做事的 基本道理



- (C)《周易》、《老子》、《莊子》,了解這些最近風行一時的學說 思想
- (D)《孫子》、《左傳》、《史記》、《漢書》等典籍,培養見識與 能力

題號	Р	Ph	Pl	Pa	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
29	76	87	64	89	83	78	72	60		23	6	5	6	12

題幹中述及「對於時局的分析大爲欽佩」,應可與選項中的「見識與能力」 相呼應。再說,閱讀史書,或學習歷史,應有其實際功用,分析時局,應爲其 中一項。這道題目,形式上略嫌複雜,實則十分簡單。考生只要讀懂題意,應 可答對。高分組的表現,令人失望。

第33題:

33. 以下是有關中國大陸繪畫藝術發展的三段描述:(甲)這期間,除了政治宣傳繪畫,其他藝術探索均處於休克狀態,畫家林風眠遭受迫害,被打為「黑畫家」,許多作品被毀,中國藝術經歷了一場空前浩劫。(乙)此時,對西方文化基本封閉,遵從共產黨領導與形式大眾化的延安革命傳統,以及學自蘇聯的社會現實主義,結合成一種美術的新正統模式。(丙)這個階段裡,國外美學、心理學、文化學和藝術學著作被廣泛介紹進來,美術理論研究也從復甦到初步繁榮,美術報刊空前增多,論爭也趨於頻繁和活躍。依其發展順序,下列選項何者正確?
(A)甲乙丙 (B)乙丙甲 (C)乙甲丙 (D)丙乙甲

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
33	45	56	33	59	51	46	39	30		23	8	5	7	9

這一題與 98 年指考的第 34 題,形式相似,但考生答題時的思考很不一樣。 98 年指考的第 34 題,要排列的是四段文字的「邏輯次序」,而此題基本上屬於

「時間次序」的排列。再加上題意有模糊不明之處,如「中國大陸繪畫」是指 哪一段時期?如果不是只限於共產黨統治下的中國大陸,(D)選項亦可成立。

第35題:

- 35. 資料甲:「管、蔡、武庚等果率淮夷而反。周公乃奉成王命,興師東征,作〈大誥〉。遂誅管叔,殺武庚,放蔡叔。」
 - 資料乙:「既破我斧,又缺我斯(方形的斧)。周公東征,四國(管蔡 商奄)是皇(同惶,恐懼)。哀(可憐)我人(兵士自稱)斯, 亦孔(很)之將(好)。」

依據上述資料,下列說明何者正確?

- (A)資料甲出自《史記》,資料乙出自《詩經》,皆述及同一歷史事件
- (B)資料甲出自《史記》,資料乙出自《詩經》,述及不同歷史事件
- (C)資料甲出自《詩經》,資料乙出自《史記》,皆述及同一歷史事件
- (D)資料甲出自《詩經》,資料乙出自《史記》,述及不同歷史事件

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
35	79	95	57	97	91	85	74	47		38	6	6	11	27

這一題與98年指考第25題頗爲相似。考生必須讀出兩段文字出自不同的 書寫體裁,一爲史書,一爲經書。再能讀懂兩段文字所述爲同一事件,即「周 公東征」,即可答對。

第36題:

36.以下是第一次世界大戰結束之後,有關英國婦女地位的兩份資料。 資料甲:一份地方報紙:「女性還不明白一件事:她們已經不太可能 再從事那些有薪資的工廠工作。先前離開家庭雜事進入工廠 的女性,如今被要求回到她們原來的鍋、盤世界中。」



資料乙:一位現代史家說:「把選舉權視為是對戰爭中女性作出貢獻 的報償,這是很簡單的解釋。實際上,仔細研究後會發現, 戰爭帶來的改變不多。那時報紙報導女性得到熱誠歡迎,但 在農場、醫院和工廠,她們都被痛恨、敵視。」

關於以上兩份資料的見解,下列敘述何者為宜?

(A)甲和乙認為戰後社會對女性的看法未改變

(B)甲和乙認為大戰改善了女性的經濟地位

(C)甲和乙認為大戰改善了女性的社會地位

(D)資料甲的觀點和資料乙的觀點互相有矛盾

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
36	72	86	57	89	80	74	67	51		29	9	6	7	16

題幹的兩段文字,觀點相同嗎?如果覺得並沒有很大的差別,把選項中有 「改善」與「矛盾」的字樣除掉,即可答對。考生略具閱讀能力,答卷時又能 專心一致,應該不難。

多重選擇題中,有兩題屬於「閱讀」型的題目。

第40題:

- 40. 史家甲:「文藝復興不只是文學或藝術的再興,也是人類智識和個人 之覺醒或再生,而為近代世界的開始。」史家乙:「文藝復興是中古 時代轉移至近代的過渡時代,它同時具備了兩個時代的特性。」有關 兩位史家對「文藝復興」的看法,下列敘述何者恰當? (A)史家甲強調文藝復興與中世紀的差異和斷裂
 - (B)史家乙強調文藝復興仍然保有中世紀的特質
 - (C)兩位史家對於文藝復興的解釋並無明顯差別
 - (D)史家甲關注歷史的延續和連結
 - (E)史家乙側重歷史的新異和變遷

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
*40	61	80	38	83	73	65	55	29	52	42	10	8	10	26

題幹文字主要是兩位史家對重要歷史時期的各自見解,考生要能辨識出他 們各自觀點的特色,也需要有從簡要的文字中看出「變遷」觀念與「延續」觀 念的能力。考生如果從未有過這種題型的練習,一時不易適應,不知從何下手。 如果做過一些類似的題目,必然覺得容易。

第38題:

38.《漢書·藝文志》著錄的醫學著作有《黃帝內經》18卷,內容多以黃 帝與人問答的形式來呈現。《黃帝內經》主張人必須「和於陰陽,調 於四時」,才不致生病。它對臟腑及經絡之功能特別注意,提出了「善 診者,察色按脈」,「經脈流行不止,環周不休」等觀念。對於此書下 列說法何者為宜?

(A)此書假借黃帝問答,可見是一本荒誕不經的偽作,不可信
(B)此書假借黃帝問答,是古代依託聖人心態下的產物,仍有其價值
(C)此書之陰陽調和和四時變化的醫學理論,已被現代的中醫所摒棄
(D)此書之臟腑、經絡學說是中醫重要基本理論,指導中醫臨床實驗
(E)此書重視陰陽四時之變化,顯然是受到陰陽五行學說的影響

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
*38	78	89	63	90	87	83	77	55	62	26	3	4	6	22

這道題目,形式上是史料閱讀題。但從題幹與選項的關係看,兩者的聯繫 並不很密切。也就是說,只要對現代中醫的情況有所瞭解,就能答對這道題目。 再說,五個選項,基本上是五個單獨的是非題。多重選擇題中,此類題型的題 目常能見到,學生的回答,仍以記憶為主。


99 年的指考試題中,我們看到了過去很少見到的題型,亦可一談。如 第4題:

4. 人類各大文明的發展,並不一致,而是有快有慢,緩慢的時候,甚至 可以譬喻為冬眠。某學者指出,十五世紀以後,這個地區逐漸從千年 的冬眠中甦醒過來,可以與某一國在哲學、數學、藝術上相競爭。請 問這個地區與某一國,各指何處? (A)歐洲、中國

(C)美洲、西班牙

(B)阿拉伯半島、印度

(D) 東亞、英國

題號	Р	Ph	Pl	Ра	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
4	59	83	35	88	73	58	44	30		48	15	15	14	14

這道題目是一個記憶題。但要記的內容,不論時間或空間,都很綿長和遼 闊,考生需要對於世界的歷史有一個大的圖像,如有,即可答對。

另外,**第21題:**

21. 中國古代科舉制度是影響深遠的考試制度,它的成立要符合一些依 準,下列敘述何者適宜?

(A)應試者自行報名,不分科目,分數及格,即可任官

(B)應試者自行報名,分科應試,中進士者可以任官

(C)應試者由地方長官推薦,不分科目,共同競爭,中試者可以任官 (D)應試者由朝廷大臣徵召,分科應試,分數及格,即可任官

題號	Р	Ph	Pl	Pa	Pb	Pc	Pd	Pe	Т	D	D1	D2	D3	D4
21	67	84	50	87	76	69	57	46		34	11	7	12	11

是關於中國史上一個重要的制度--科舉制度--的題目,考生需要對這個制度 的幾項重點有清楚的記憶方能答對。只是這道題目的選項設計有瑕疵,考生只 要選有「進士」字樣的選項,即可答對。

我的意思是,「記憶」類的試題並不是不好,也不是應該廢去,而是到了有 所調整的時刻了。今天,我們看到學生對歷史這門學科的認知就是「記憶的科 目」,應付考試的方法就是死記硬背,只是背過即忘。而且多年以來,歷史試題 偏難,致使學生學習興趣低落。面對此一情況,把歷史考試的試題,從「記憶」 類中,部分轉變為「閱讀」類。由於解題容易,考生得分必能提高。對於學生 學習興趣,亦將有提升的作用。重要考試多用「閱讀」型的試題,應是針對目 前學生在學習歷史時的錯誤觀念與作法,提出的一項有效改進策略。再說,這 樣以「閱讀」為主的「新」題型,也能夠符合歷史教學的考試「理論」,也可以 說是把歷史教學理論加以落實的有效作法。至於「舊」的「記憶」類試題與「新」 的「閱讀」類試題,應該以怎樣的比例搭配,或者在選擇題中還要開發其他類 型的試題,都是我們需要進一步研究的問題。

肆、試題命製的建議

命製試題是我們教師的工作之一,我們如何命製可以引導學生進入歷史領 域,對歷史知識的結構、方法有所接觸的選擇題呢?我認為,首先要想到,我 們在什麼地方可以找到命題的素材。教科書中是沒有的,因爲教科書篇幅有限, 可用的資料太少;再說,教科書限於體例,也不可能有足夠的論證。因此,命 題的素材就在資料與論證豐富而又集中的地方。一般的學術著作,資料與論證 皆有,但各有主題,大多過於狹窄,難以取材;學者的一得之見,往往過於偏 頗,尤不可用。那麼,我們要在那裡找到可用的素材呢?個人的淺見,大學者 的一般性著作,最爲適宜。大學者著述,資料豐富之外,論證尤其嚴謹;一般 性著作,範圍廣大,可資運用者,隨處皆有,最稱方便。舉例言之,中國史方 面,錢穆的《國史大綱》與郭廷以的《近代中國史綱》就是最可以仔細閱讀, 並從中取材,命製試題的兩部「一般性史學著作」。

茲舉錢著《國史大綱》為例,試製三題,以供參考。必須聲明的是,我這 裡的題目,主要是從引導學生進行思考的角度來命製,而不是為了檢測學生的 學習成果。我相信學生只有多做閱讀思考型的歷史試題,他的歷史思考能力才 得以提升。

一、《國史大綱》頁 103,錢穆說:儒家可以稱為「良心教」,墨家可以稱 為「苦行教」。良心與苦行,皆代表中國民族精神之一部。惟苦行究極必本於良 心,若專本諸天志,則其事不可久。而良心則不必限於苦行。故儒可以兼墨, 墨不足以代儒。我們可以取這段話的前面大部分作爲題幹,留下最後一句話, 置於選項。題目最前,寫上諸如「一位著名學說」或直接寫「錢穆說」,引文作 爲題幹的主體,然後問:「這段話,可以得出怎樣的論斷?」或「這段話可以做 怎樣的概括?」選項設計為:

- (A) 儒、墨雨家學說並無不同, 皆源於愛;
- (B) 儒、墨兩家學說完全不同, 全無交集;
- (C)墨家可以兼儒家,儒家不足以代墨家;
- (D) 儒家可以兼墨家,墨家不足以代儒家。

這道題目的好處有:一、問的是一個大問題,儒家與墨家的關係,學生理 應知悉。二、錢穆的對比很有意思,「良心教」與「苦行教」是一個有趣的譬喻, 而論述也有其理路,很可以讓學生閱讀之時,加以思考。不過,這題亦有一個 明顯的缺點,就是題幹與選項的關係並不密切,學生只要對儒、墨關係有所認 知,不須閱讀題幹即可答對。如果做爲學生的閱讀與思考練習,這個缺點應該 不是很嚴重。

二、《國史大綱》頁 161,論及漢代外戚的興起,我們可以命製以下這樣一 道題目:

西漢初年,大封同姓。吳、楚七國亂後,宗室地位日見低落。錢穆先生的 解釋是:「宗室只宜封建,不宜輔政,以其地近而勢尊。封建政治既不能復活, 則宗室地位自難再興。」請問:錢先生的意思是:

(A) 宗室宜封建,以其地近勢逼,宗室多在京城附近,威脅很大;

(B) 宗室宜封建,以其地近勢逼,宗室多屬皇帝兄弟,威脅很大;

(C) 宗室不宜輔政,以其地近勢逼,宗室多在京城附近,威脅很大;

(D) 宗室不宜輔政,以其地近勢逼,宗室多屬皇帝兄弟,威脅很大。

這道題目,只是看學生能否讀懂「地近勢逼」這四個字的意思,並不困難。 但它談了一個大問題,七國亂後,封建政治再也不能復興情況下,宗室處於怎 樣的地位。學生讀了題幹與選項,其實就是對於中國史上的一個重要現象,做 了一次思考的練習。

三、《國史大綱》頁 283,關於北魏孝文帝遷都華化的資料,可命製一題, 題幹為:

北魏孝文帝說:「北人每言北俗質魯,何由知書?朕聞之,深用憮然。今知書者甚眾,豈皆聖人,顧學與不學耳。朕為天子,何必居中原?正欲卿等子孫, 漸染美俗,聞見廣博。若永居恆北,復值不好文之主,不免面牆耳。」

問題可以從多面思考。如:

1、北魏孝文帝的這段話,談及什麼事情?

(A)變法 (B)征伐 (C)興學 (D)遷都2、北魏孝文帝主張遷都,是從什麼方面考慮?

(A)政治 (B)社會 (C)經濟 (D)文化3、北魏孝文帝的這段話,是對什麼人所說?

(A)朝廷中的漢人官員 (B)抵達洛陽的胡人將士

(C)積極支持遷都的大臣 (D)對遷都感疑慮的大臣

北魏孝文帝的華化遷都,是魏晉南北朝的歷史中,學生少數熟悉的史事之 一,我們問第一個問題,學生立即可以運用記憶庫中的資料回答。問第二個問



題,他們就要想一想,這段話究竟是與何者有關?需要想一想,做一點思考練 習的。第三個問題,似乎比較困難,學生必須從文字中讀出說話人心中的想法。 如果學生看不懂像是「深用憮然」(深深感到不以爲然)或「永居恆北」(一直 居住在北方),就不大容易讀得懂。我們的處理方法有二,一是把選項的距離拉 大,就像上題的作法;二是把不容易懂的文字加上注釋,也就是把上文括號中 的文字加入題幹中。我們設計學生練習思考的題目,可以把難度儘可能降低, 關鍵在於一定要學生想一想,用用腦筋。另外,我們還可以利用這段資料設計 問題,只取引文作爲題幹,問:「這段文字,是關於歷史上的什麼事情?」選項 爲:

(A) 漢武帝獨尊儒術(B) 曹操率大軍南下

(C) 苻堅大舉南征 (D) 北魏孝文帝遷都

也許,學生比較會做這樣的題目,因為他們對於這樣的題型較為熟悉,做 過許多這樣型式的題目。然而,這道題目可以檢測出學生的閱讀能力,仍然是 可以一用的。

總之,錢穆的《國史大綱》中,選用的資料都很精要,論斷亦很精審,我 們仔細閱讀,不能不佩服錢先生學識之博與思考之深。書中的敍事無不有根有 據,而識見與評論尤其細密嚴謹,我們在課堂教學中,可以學習、取材的地方 甚多,命製試題,只是其中一項而已。

參考文獻

大學入學考試中心(2009)。98年指定科目考試歷史考科試題及答對率與鑑別指數表。

取自:http://www.ceec.edu.tw/AppointExam/AppointExamProfile.htm 大學入學考試中心(2010)。99年指定科目考試歷史考科試題及答對率與鑑別指數表。

取自:http://www.ceec.edu.tw/AppointExam/AppointExamProfile.htm 何兆武(1995)。**歷史與歷史學**。香港:牛津大學出版社。

錢穆(1995)。國史大綱。台北:台灣商務出版社。