

閱讀選擇題第一篇

壹、前言

長久以來，國民小學的社會科教學，面臨著不易克服的困境，這些困境，可以由以下兩點來說明：

一、地位上最不受重視的學科：社會科教學，一向不很受重視，教學品質，相對的就難以提升。這種現象，從幾點事實可以看見：1. 學校在安排任課教師時，社會科的人選往往最後考慮，認知上就以為社會科人人可教，不需任何條件或素養；2. 難得為社會科設置專科教室，即使設了，也是虛有其表，欠缺基本的設施；3. 教具的充實，往往也是國語、數學、自然優先採購，社會科教具往往被忽略，甚至有連藝能科都比不上的情況。

二、教學上最偏離目標的學科：社會科的性質，除了知識的學習以外，是最具陶冶情意、增進技能功能的學科；教學目標的設定，也著重情意與技能。然而，由於以下的不利因素，卻造成了教學的偏離目標：1. 教學多用近似「讀書」的方式，未能「活動化」；2. 學習只重知識的記憶，忽略了社會科最重視的「概念」與「通則」的形成；3. 評量以紙筆測驗為主，忽略平時的評量；以給分為例，尚少有技能和情意的評量；忽略了「評量多元化」的要求。這些情況，在在顯示出，社會科的教學，難以達成其該有的教學目標。

在這種情況之下，社會科的教學，本已有所不足。當時勢所趨，在教育改革、教學求新的需求下，欲實施「非紙筆測驗」的教學評量，也就遇上一些問題。這些問題，可以從教師、學童、家長三個角度，看出端倪：

一、教師方面：由於過去慣於以習作、紙筆測驗作為評量方式。在毫無頭緒的情況下，採用哪種方式頗耗思量；其次，如何讓學童了解評量過程與要求標準，也相當費口舌；最後雖評出分數，但因欠缺具體工具，以致受到學童、家長質疑，解說起來勞心費神。

二、學童方面：由於沒有作答經驗，一時無所是從，到底老師要求的是什麼，不易把握；為了完成老師的要求，短期間付出更多的時間和精神，反而形成了另一種壓力。

三、家長方面：為求好分數，在學童進行作業過程中，協助預備，甚且急著代為捉刀。一個以學童學習成果為對象的評量，變成家長角力的焦點，盡失評量的意義。

然而，這些表徵的產生，基本因素就在於對實施「非紙筆測驗」的誤認：

一、誤以為「非紙筆測驗」能減輕學童負擔，實際上，在無所是從、不知如何著手的壓力下，學童的負擔，反而增加。

二、在對「非紙筆測驗」評量方式與實施程序，不甚了解的情況下，為了分數的評定，只給學童很短的時間，卻要求去完成「非紙筆測驗」的題目，漠視「非紙筆測驗」評量的精義，盡失其價值。

三、以平時的「紙筆測驗」，做為定期考查「非紙筆測驗」成績的一部分，實質上還是「紙筆測驗」，只是換湯不換藥。

四、高估學童能力，只要求個別式的「非紙筆測驗」，結果是程度好的學童勉力完成；程度平常的學童，由於不知如何下手，往往敷衍了事。這種情況下，難以讓學童的能力，

有適當的展現。

改進教學評量，運用多種方式來進行，是今天教學上不可免的趨勢。倡導「非紙筆測驗」教學評量的意義，就在：揚棄紙筆測驗為主的評量觀念，為教師專業自主的發揮，奠立基礎。然而面對存在的問題，最重要的還是在：釐清「非紙筆測驗」教學評量的內涵，提供具體的參考模式，喚起教師根據教學評量原理，落實「非紙筆測驗」教學評量的實施。

貳、社會科非紙筆測驗教學評量的定位

「非紙筆測驗」教學評量，在社會科中如何加以定位？應該是探討社會科「非紙筆測驗」教學評量的前提。以下將從社會科課程內涵與特性、以及社會科特別重視技能和情意的培養這兩點，來加以探討。

一、檢視社會科課程內涵與特性：

- (一)從社會科課程理想內涵來看，內容豐富，包括：1. 生活領域：以個人、家庭、學校、地方、國家、世界為範圍的知識；2. 個體行為：包含品德修養與生活習慣兩方面；3. 社會行為：包含人際關係與群己活動、物質生活和社會經濟、文化生活和社會福祉等三方面；4. 學習基本能力：包含閱讀、發表、欣賞、解決問題、蒐集及運用資料等。
- (二)從社會科課程學科領域來看，範圍寬廣，包括了：歷史學、地理學、政治學、經濟學、社會學、心理學。
- (三)從國民小學社會科的特性來看，歸納為：1. 具鄉土性的學科：由於學習者的身心背景的條件，教學內容著重「鄉土性」的素材；2. 具時事性的學科：隨著時事的移易，教學材料就必須不斷的推陳出新；3. 最多元反應的學科：社會現象是不確定的，也沒有固定的解決方式，或放諸四海而皆準的答案；4. 最具價值的學科：社會科建立人類社會基本的價值體系，是最具價值的學科；5. 富挑戰性的學科：教師必須有豐富的知識，始能勝任愉快，對教師而言，是隨時需要面對挑戰的一個科目。

經過一番檢視以後，可以發現：社會科內容豐富、範圍廣闊、特性顯著，在進行教學評量時，非用多種方式，實難有效評鑑學習的成果。

二、社會科特別重視技能和情意的培養：

考查國民小學社會科技能教學與情意教學的內涵，可以發現社會科教學，在達成這兩層面的目標的功能和重要：

- (一)國民小學社會科技能教學內涵：1. 生活規範的實踐；2. 學習活動的參與；3. 器物使用與珍惜；4. 閱讀與繪製圖表；5. 蒐集與運用資料；6. 社會問題探討與解決。
- (二)國民小學社會科情意教學內涵：1. 盡己能力和責任；2. 尊親、友愛與和睦；3. 團體合作和和諧；4. 尊重行業和服務；5. 公德、同情與惜物；6. 社會道義與參與；7. 國家觀念與民族情感；8. 世界觀與大同理想；9. 民主風度與政治參與；10. 愛校、

愛家、愛鄉土、愛國家。

從這些技能與情意的教學內涵來看，社會科教學評量的進行，單藉紙筆測驗，難以竟其功。運用多種的評量方式，配合目標層次來進行，才是踏實的做法。

根據簡茂發的看法，教學評量原則有：1. 應是多方面的一除了知識技能之外，應包括學童的品行、人際關係，學習態度、興趣、方法和習慣等方面；2. 必須採用多種方式；3. 應注意學童平時的學習活動。又歐用生認為，社會科教學評量可遵循原則有：1. 應視為教學的一部分；2. 併用多種評量技巧；3. 利用各種場合評量。李聰超綜合各家意見，也提出了社會科教學評量，應該：1. 揚棄以紙筆測驗為主的觀念；2. 以協助兒童的學習與成長為依歸；3. 兼用各種評量方式；4. 重視兒童學習中的表現，強調平時的評量。這些看法，都在指出了一個道理：社會科教學評量應以「非紙筆測驗」為主。同時也符合社會科課程內涵與特性，因應社會科特別重視技能和情意的培養的事實。

參、社會科非紙筆測驗評量方式的探討

教育部公布的國民小學成績考查辦法中，就成績考查的方式，總共列了十一項。其中屬「非紙筆測驗」的方式，有：鑑賞、晤談、報告、表演、實作、資料蒐集、設計製作、作業、實踐及其他。茲就社會科教學實施中，較適合運用之方式，分別陳述如下：

一、報告：意指是「就學童閱讀、觀察、實驗、調查等所得結果之書面或口頭報告。」在社會科教學評量運用時，應該注意：1. 報告的主題，應是教學的重要內容；2. 報告的主題，應符合教學目標；3. 應給兒童充分之反應時間，並確切的告知；4. 使用時間、參考資料、報告方式等事項，都應明確告知；5. 界定評量標準，並向學童明確提示；6. 可以讓其他學童，共同參與評量。

二、表演：意指是「就學童之表演活動加以考查。」在社會科教學中實施時，宜把握下面的幾項原則：1. 表演內容或主題，由兒童共同討論決定；2. 個別表演或團體表演均可運用，但以小組較佳；3. 以質為主要依據，以把握主題之程度為首要；4. 應訂定評分標準，以為依循，並事先告知；5. 應給與充分的時間，使能全力演出。

三、實作：意指是「就學童之實際操作及解決問題等行為表現加以考查。」以下是給社會科教師的建議：1. 實作的項目，應是能具體操作的內容；2. 擬具操作的要項，逐項評定，並事先告知學童；3. 協助學童，準備必要的器材；4. 評量標準事先界定，並告知學童；5. 操作過程儘量納入評量項目，而非只重成品或結果。

四、資料蒐集：意指是：「就學童對資料之蒐集、整理、分析及應用等活動加以考查。」在社會科教學中運用時，宜注意：1. 應有一定的主題或範圍，並符合教學目標；2. 注重所得資料的歸類、整理，而非僅剪貼；3. 資料的質為主，量為輔；4. 應給予足夠的時間，使能充分搜羅；5. 小組進行時，應考慮每個人努力的程度；6. 搜集過程中，正確的運用資料的態度，應列入評量。

五、設計製作：意指是「就學童之創作過程及實際表現加以考查。」社會科教師從事設計製作的考查時，應：1. 提示主題或具體內容；2. 提示所需工具，指導學童準備；3. 過程與結果應同時重視；4. 擬具評量的要項，逐項評定。

六、實踐：意指是「就學童之日常行為表現加以考查。」應：1. 設計具體實踐項目，並明確告之要求標準；2. 日常生活中隨時進行，以求真實；3. 運用自我評量、同學互評等方式，以供參考；4. 採用評核表記錄，以便整理。

如果從目標層次與運用方式的配合來看，哪些方式適用於哪一目標層次，是教師進行非紙筆測驗教學評量時，所感到困擾的問題。茲就其適用情況，說明如下：

一、屬於認知目標的：晤談、報告、問答討論、資料蒐集、作業。

二、屬於情意目標的：鑑賞、表演、實踐。

三、屬於技能目標的：表演、實作、資料蒐集、設計製作。

綜言之，從上面的探討，應該可以對「非紙筆測驗」評量方式的運用，做以下的歸納：

一、非紙筆測驗的教學評量，應融入教學過程中，與教學活動實際上是分不開的。

二、選用合宜的方式交互運用，不能單用一種方式進行評量，就代表學童學習的成果。

三、評量不能單著眼在給分，應該是透過評量的進行，給予立即的回饋，以增進學習的效果。

四、教師善於設計合用的評量表，以便隨時登錄運用，使評量更為客觀、公正。

肆、社會科非紙筆測驗評量的基礎

社會科「非紙筆測驗」評量如何進行？教師必有的基本體認是什麼？以下就實施原則、要點以及評分方式等，分別加以說明：

一、實施原則：1. 依據教材的性質，選用合適的評量方式：如社會現象用資料蒐集、歷史事實用表演、社區問題用書面報告、閱讀地圖用實作等；2. 考量學童背景：學童的能力以及可運用的資源和設施，都是進行教學評量時，所應考量的；3. 評量進行中營造競爭氣氛，使願意盡力完成工作：創造激勵憤發向上的環境，讓學童產生「追求更好」的意願；4. 兼顧深度和廣度，讓學童按自己能力發揮：使程度高的學童，能有更精到的表現，程度平常的學童，做基礎的反應；5. 以分組作業為原則，用個別分數來呈現：小學學童，受經驗、能力的限制，較難獨力完成作業，以分組方式來設計，是必要的。

二、實施要點：1. 事前的周詳計畫：先了解教材及教學目標、擬訂評量的實施計畫，將主題、成果呈現方式、繳交日期、評分項目、評分標準、參考資料等，都列入評量計畫中，告知學童。2. 適切的進行分組：每組人數以六人為度，有利彼此的互動；各組的實力應該相等，以激發努力求好的鬥志；每次評量後，都應重新分組，促進人際關係的調適；鼓勵能力較差的學童，使他們在小組中，能盡力表現；開導小組每位成員，發揮接納別人的愛心，避免排斥能力較差的學童。3. 給予小組討論的機會：在各組進行搜集、準備的階

段，教師宜利用上課時間，留給小組就題目進行討論，並藉以了解各組的進度，督促其在限定時間內完成；了解各組的困難，作必要的提示；了解學童個別參與的狀況，以適時鼓舞。

三、評分方式：在以小組合作為基礎的前題下，分數的評定，宜分為「小組分數」和「互評分數」進行，再結合為「個別分數」。其評分的要點如下：

1. 小組分數：由教師依事前既定項目及標準，進行各組評分，將評分結果，轉換成百分法計分，是為小組分數。
2. 互評分數：由小組成員互相評分，但本人不做自評。將小組成員評分結果，合計為總分。依個別所得總分，排列高低順序。
3. 個別分數：教師先行決定等級之間間距，以 2~3 分為度。以小組分數為基準，做為排序居中一人的成績。由中間依互評分數排序，向兩頭加減間距，計算各組成員成績。

舉例來說，假設小組分數，經教師評分為 85 分，決定差距為 3 分。此時，若經小組成員互評後，學童甲排序 3、學童乙排序 5、學童丙排序 2、學童丁排序 1、學童戊排序 4，則該小組學童的個別成績為：排序居中(排序 3)的學童甲得 85 分、排序 2 的學童丙得 88 分、排序 4 的學童戊得 82 分，然後排序 1 的學童丁得 91 分、排序 5 的學童乙得 79 分。

伍、社會科非紙筆測驗評量的模式

現就社會科非紙筆測驗評量的模式以「資料搜集」、「表演」為例分別說明：

一、資料搜集的評量：首先擬定計畫，然後訂定教師的小組分數評量表及建立小組內學童互評表，最後進行評分工作，茲以「搜集萬華區古蹟的相關資料」為題說明要點如下

1. 擬定評量計畫：

- (1) 年級單元：社會科三年級下學期第一單元「認識我們的家鄉」。
- (2) 評量題目：搜集萬華區古蹟的相關資料。
- (3) 成果呈現：書面資料，以 A4 白報紙張貼，並加封面，以十張為限。圖片不拘，其他(訪問錄音)，自由運用。
- (4) 繳交日期：四月十日，放學以前。
- (5) 參考資料：臺北市政府民政局出版—臺北市的古蹟；臺北市政府新聞處出版—臺北市古街之旅；遠流出版公司出版—臺北歷史散步；臺北鄉情叢書國小篇第二冊—臺北的故事；各古蹟管理單位出版的簡介。

2. 訂定教師小組分數評量表如下

評量項目	0	1	2	3	4	5
1. 搜集範圍的完整	<input type="checkbox"/>					
2. 搜集材料切合主題	<input type="checkbox"/>					

評量項目	0	1	2	3	4	5
3. 搜集材料分別整理	<input type="checkbox"/>					
4. 編排有次序	<input type="checkbox"/>					
5. 提要切合內容簡扼	<input type="checkbox"/>					
6. 提要文字生動流利	<input type="checkbox"/>					
7. 資料來源詳細註記	<input type="checkbox"/>					
8. 圖片資料有價值	<input type="checkbox"/>					
9. 運用各種方法搜集	<input type="checkbox"/>					
10. 小組分工合作	<input type="checkbox"/>					

此表需事前訂定，每項成績為 0~5 分，例如三年一班第一小組從第 1 項~第 10 項分別得分 4, 5, 4, 4, 3, 4, 5, 3, 4, 3 時，第一小組的分數為：

$$\frac{4+5+4+4+3+4+5+3+4+3}{10 \times 5} \times 100 = 78$$

3. 建立小組內學童互評表：此表讓小組內每位學童評分，最低 1 分，最高 5 分。例如三年一班第一小組評分者陳明禮對組內的評分如下：

評量項目	漢民	正平	筱娟	一萍	怡嘉
1. 搜集盡責	4	3	5	4	5
2. 熱烈討論	4	3	4	5	4
3. 提供器材	5	4	4	3	3
4. 助人搜集	3	3	4	4	3
5. 幫助摘要	4	3	3	4	5
6. 和人合作	3	4	2	5	4
小計	23	20	22	25	24

然後將每位互評分數經加總後，按總分將每位組內學童加以排列順序如下：

姓名	漢民	正平	筱娟	一萍	怡嘉	明禮
明禮	23	20	22	25	24	
怡嘉	18	17	20	22		20
一萍	22	18	19		24	23
筱娟	21	17		23	20	22
正平	23		22	22	23	21
漢民		19	21	22	22	18
總計	103	91	104	114	113	104
排序	3	6	4	1	2	4

4. 個人得分計算：三年一班第一小組的小組分數是 78 分，如果事前決定差距為 3 分，則個別分數為

姓名	漢民	正平	筱娟	一萍	怡嘉	明禮
排序	3	6	4	1	2	4
得分	78	72	75	84	81	75

二、表演的評量：首先擬定計畫，然後訂定教師的小組分數評量表及建立小組內學童互評表，最後進行評分工作，茲以「民族融合故事的編劇演出」為題，說明要點如下

1. 擬定評量計畫：

- (1) 年級單元：五年級上學期第九冊第二單元「中華民族的融合」
- (2) 評量題目：民族融合故事的編劇演出。
- (3) 成果呈現：劇本編寫以 A4 紙兩張為準。
- (4) 表演方式：現場演出十分鐘。
- (5) 表演日期：十月十五日第二、三節
- (6) 參考資料： 中華兒童叢書－中國的服飾； 烏恩著、中華書局出版－百川歸海；
 光復書局出版－中國歷史圖說； 小魯出版社出版－中國歷史 365 小百科

2. 訂定教師小組分數評量表如下

評量項目	0	1	2	3	4	5
1. 劇本切合主題	<input type="checkbox"/>					
2. 劇本生動有趣	<input type="checkbox"/>					
3. 劇情合邏輯	<input type="checkbox"/>					
4. 根據史實編寫	<input type="checkbox"/>					
5. 角色把握個性	<input type="checkbox"/>					
6. 道具盡力製作	<input type="checkbox"/>					
7. 表演賣力	<input type="checkbox"/>					
8. 小組合作無間	<input type="checkbox"/>					
9. 演出井然有序	<input type="checkbox"/>					
10. 時間把握適當	<input type="checkbox"/>					
11. 參考資料充分	<input type="checkbox"/>					
12. 演後整理確實	<input type="checkbox"/>					

此表需事前訂定，每項成績為 0~5 分，例如五年二班第二小組從第 1 項~第 12 項分別得分 5, 4, 4, 3, 3, 4, 4, 4, 5, 4, 4, 5 時，第二小組的分數為：

$$\frac{5+4+4+3+3+4+4+4+5+4+4+5}{12 \times 5} \times 100 = 82$$

3. 建立小組內學童互評表：此表讓小組內每位學童評分，最低1分，最高5分。例如五年二班第二小組評分者李淨對組內的評分如下：

評量項目	曉明	康靖	浩學	維真	美慧
1. 搜集盡責	5	3	2	4	4
2. 熱烈討論	4	3	3	4	5
3. 參編劇本	5	2	3	3	4
4. 樂接角色	5	3	4	4	5
5. 排練合作	5	3	3	3	4
6. 熱心道具	3	4	3	5	4
7. 全力演出	5	3	2	4	5
8. 樂於助演	4	2	3	3	3
9. 收拾整理	3	5	4	4	3
小計	39	28	27	34	37

然後將每位互評分數經加總後，按總分將每位組內學童加以排列順序如下：

姓名	美慧	維真	李淨	浩學	康靖	曉明
美慧		36	33	29	30	40
維真	33		29	25	26	36
李淨	37	34		27	28	39
浩學	36	33	31		31	38
康靖	37	35	31	33		41
曉明	34	30	30	29	27	
總計	177	168	154	143	142	194
排序	2	3	4	5	6	1

4. 個人得分計算：五年二班第二小組的小組分數是82分，如果事前決定差距為2分，則個別分數為

姓名	美慧	維真	李淨	浩學	康靖	曉明
排序	2	3	4	5	6	1
得分	86	84	82	80	78	88

陸、結語

教學評量的改進，是因應教育改革、教學發展的必然趨勢，身為教師，應該不能自外於職責所在，在「非紙筆測驗」教學評量，盡一分心力，以開創教學的新局面。

社會科就學科內涵和特性，以及學科重視情意、技能的培養來看，教學評量，應該以「非紙筆測驗」的方式為主，如此，當有助於社會科教學水準的提升。然而，由於社會科的課程實施，長久的不受重視，加上「非紙筆測驗」一直未能在教學評量中，佔一席之地，驟然實施、推廣，面臨的是教師、學童都不知所措的情況，有待教育工作同仁的探討、解決。

本文試就「非紙筆測驗」評量方式加以探討，再從「非紙筆測驗」的基礎——評量表的運用、「非紙筆測驗」評量的基礎及模式加以說明，期望提供國民小學教師社會科「非紙筆測驗」評量的參考。當然，這只是「初探」，寄望國民小學教師，不斷的探討、交換心得，建立非紙筆測驗評量的可行模式，讓「非紙筆測驗」的評量，落實在社會科的教學實施中。

請依上文所述，選擇最適當的選項，回答下列問題：

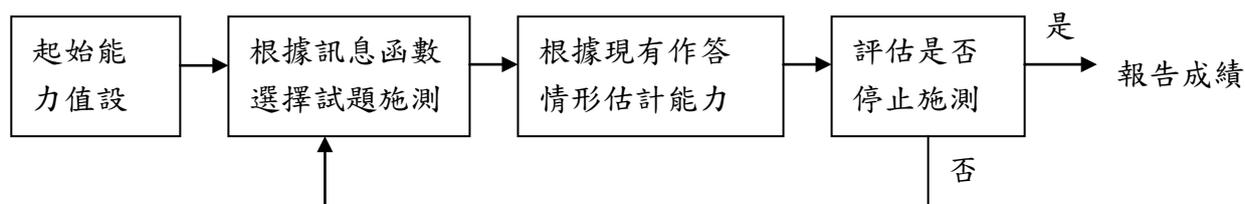
1. 依本文所述，下列哪一個選項是現今社會科很難達成教學目標的原因之一？
 - (A) 社會科太重視教學活動化
 - (B) 社會科太重視評量多樣化
 - (C) 社會科太重視概念通則化
 - (D) 社會科太重視知識記憶化
2. 對「解決問題」的行為表現加以考查，是屬於下列哪種「非紙筆測驗」的方式？
 - (A) 實作
 - (B) 實踐
 - (C) 作業
 - (D) 晤談
3. 依本文所述，下列關於社會科「紙筆測驗」與「非紙筆測驗」的比較，何者最為正確？
 - (A) 「非紙筆測驗」較可以降低學童的壓力負擔。
 - (B) 「非紙筆測驗」較可以減少學童的測驗時間。
 - (C) 「非紙筆測驗」較容易施行學童的情意評量。
 - (D) 「非紙筆測驗」較容易忽視學童的認知回饋。
4. 依本文所述，下列 5 個項目的社會科測驗與評量中，哪幾項為「非紙筆測驗」？
作業評量；表演評量；基本學力評量；設計製作測驗；電腦適性測驗。
 - (A) 僅是
 - (B) 僅是
 - (C) 僅不是
 - (D) 僅不是
5. 依本文社會科「表演」的評量模式，小組內的互評表項目中，最多屬於下列哪項目標？
 - (A) 認知目標
 - (B) 情意目標
 - (C) 技能目標
 - (D) 綜合目標

6. 依本文社會科「資料搜集」的評量模式，教師訂定小組分數評量表的項目看來，最多是屬於下列哪個選項的目標？
- (A) 認知目標、情意目標
 - (B) 情意目標、技能目標
 - (C) 技能目標、認知目標
 - (D) 綜合目標
7. 依本文所述，下列 5 個項目的社會科「非紙筆測驗」方式是屬於情意為目標的評量？
鑑賞；晤談；報告；實踐；實作。
- (A) 僅
 - (B) 僅
 - (C) 僅
 - (D) 僅
8. 依本文所述，假設進行一項社會科「非紙筆測驗」，若其中的一個小組的評量表項目所得的分數分別為 3, 4, 5, 2, 0, 4 時，試問該小組分數為何？
- (A) 60
 - (B) 62
 - (C) 64
 - (D) 66
9. 依本文所述，假設進行一項社會科「非紙筆測驗」，若其中的一個小組成員有 5 名學童，若經教師評定小組分數為 81 分且設定名次間距為 3 分，試問該小組的第一名學童的個別分數為何？
- (A) 84
 - (B) 85
 - (C) 86
 - (D) 87
10. 從測驗與評量理論觀點來看，關於本文所舉的「資料搜集」、「表演」的兩項評量模式，下列何者最有可能被質疑？
- (A) 信效度
 - (B) 難易度
 - (C) 鑑別度
 - (D) 猜測度

閱讀選擇題第二篇

訊息函數(information function)在電腦適性測驗中，扮演一個重要的角色，在電腦適性測驗的進行中，藉由試題訊息函數，挑選適合於受測者的試題，使得試題對受測者為難易適中，並求得最大測驗訊息函數值，使其能力值估計的標準誤為最小，達到提高能力值估計的精確度，並且能提高測驗的效率。

在一般而言，電腦適性測驗的進行步驟為



不同的試題反應模式下，有不同的訊息函數型式，以下主要探討二元計分試題反應模式的訊息函數。試題訊息函數(item information function)，作為測驗挑選試題的主要參考依據。則試題訊息函數定義如下：

$$I_j(\theta) = \frac{[P'_j(\theta)]^2}{P_j(\theta)Q_j(\theta)} \quad (1)$$

其中 $I_j(\theta)$ 為能力值 θ 在試題 j 的試題訊息函數， $P_j(\theta)$ 為能力值 θ 在試題 j 的答對機率， $Q_j(\theta)$ 為能力值 θ 在試題 j 的答錯機率， $P'_j(\theta)$ 為能力值 θ 的試題反應函數的導數。

若試題估計是使用三參數 logistic 模式，則試題反應函數定義如下：

$$P_j(\theta) = c_j + (1 - c_j) \frac{1}{1 + \exp(-d \cdot a_j(\theta - b_j))} \quad (2)$$

其中， $P_j(\theta)$ 為能力值 θ 在試題 j 的試題反應函數， b_j 為試題 j 的難易度， a_j 為試題 j 的鑑別度， c_j 為試題 j 的猜測度。

某試題 j 在能力值 m_j 時，試題訊息 $I_j(m_j)$ 為最大，則 m_j 的公式如下：

$$m_j = b_j + \frac{1}{d \cdot a_j} \ln\left(\frac{1 + \sqrt{1 + 8c_j}}{2}\right) \quad (3)$$

從(3)可知，如使用三參數 logistic 模式，在猜測度 $c_j \neq 0$ ，則受到鑑別度 a_j 與猜測度 c_j 的影響，在能力值 θ 對以 b_j 為準而產生一個偏移量 Δm_j 。則 Δm_j 的公式如下：

$$\Delta m_j = \frac{1}{d \cdot a_j} \ln\left(\frac{1 + \sqrt{1 + 8c_j}}{2}\right) \quad (4)$$

若能力值 θ 在 m_j 點上，可得試題訊息最大值為 $I_j(m_j)$ ，則 $I_j(m_j)$ 的公式如下：

$$I_j(m_j) = \left(\frac{d \cdot a_j}{2}\right)^2 \cdot f(c_j) \quad (5)$$

在(5)中 $f(c_j)$ 的計算方式如下：

$$f(c_j) = 4(1 - c_j) \cdot \frac{2}{(1 + 2c_j + \sqrt{1 + 8c_j})} \cdot \frac{(1 + \sqrt{1 + 8c_j})^2}{(3 + \sqrt{1 + 8c_j})^2} \quad (6)$$

請依上文所述回答下列問題：

- 在公式(2)中，一般假設 d 為常數，其功用為何？
 - 為了使試題特徵函數更趨近於常態肩形模式(normal ogive model)。
 - 避免 $\exp()$ 函數值過大，而造成 $P_j(\theta)$ 不精確。
 - 修正試題的鑑別度參數，使 $P_j(\theta)$ 更為精確。
 - 修正試題的猜測度參數，使 $P_j(\theta)$ 更為精確

2. 承上題， d 通常預設值為何？

- $\log_e 2$
- 0.5
- 1.702
- π

3. 在三參數 logistic 模式下，試題訊息函數為何？

$$(A) I_j(\theta) = \frac{da_j(1 - c_j)}{[c_j + \exp(da_j(\theta - b_j))] \cdot [1 + \exp(-da_j(\theta - b_j))]^2}$$

$$(B) I_j(\theta) = \frac{d^2 a_j^2 (1 - c_j)}{[c_j + \exp(da_j(\theta - b_j))] \cdot [1 + \exp(-da_j(\theta - b_j))]^2}$$

$$(C) I_j(\theta) = \frac{d^2 a_j^2 (1 - c_j)}{[c_j + \exp(da_j(\theta - b_j))] \cdot [1 + \exp(-da_j(\theta - b_j))]}$$

$$(D) I_j(\theta) = \frac{d^2 a_j^2 (1 - c_j)}{[c_j + \exp(da_j(\theta - b_j))]^2 \cdot [1 + \exp(-da_j(\theta - b_j))]^2}$$

4. 從公式(3)可知，如使用單參數或二參數 logistic 模式時，則試題訊息最大值會發生在能力值為何？

- 鑑別度 a_j
- 難易度 b_j
- 猜測度 c_j
- 常數 d

5. 從公式(5)、(6)中，可以得出試題訊息的大小，主要受到哪些因素決定？
- (A) 鑑別度 a_j ，難易度 b_j ，猜測度 c_j ，常數 d
- (B) 鑑別度 a_j ，難易度 b_j ，猜測度 c_j
- (C) 鑑別度 a_j ，難易度 b_j
- (D) 鑑別度 a_j ，猜測度 c_j
6. 承上題，試題訊息最大值的位置，主要因素由何者決定？
- (A) 試題鑑別度 a_j
- (B) 試題難易度 b_j
- (C) 試題猜測度 c_j
- (D) 常數 d
7. 從公式(5)、(6)中得知，試題訊息的大小和試題猜測度 c_j 的關係為何？
- (A) 約與試題猜測度 c_j 成正比
- (B) 約與試題猜測度 c_j 的平方成正比
- (C) 約與試題猜測度 c_j 成反比
- (D) 約與試題猜測度 c_j 的平方成反比
8. 一般而言，電腦適性測驗的受試者初始能力值設定方式為下列何者？
- (A) 0
- (B) 介於最大能力值至最小能力值間隨機選取
- (C) 介於 0 至最小能力值隨機選取
- (D) 介於 0 至最大能力值間隨機選取
9. 假設現有 5 試題，其參數如下表所示，何者最可能被選擇為第一題進行施測？

Item	b_i	a_i	c_i
1	1.20	0.50	0.15
2	0.50	1.00	0.07
3	0.00	1.50	0.05
4	-0.50	1.00	0.09
5	-1.20	0.50	0.12

- (A) Item 1
- (B) Item 2
- (C) Item 3
- (D) Item 4
10. 下列何者為最常用的測驗訊息函數(test information function, $I(\theta)$)？
- (A) $I(\theta) = \sum_j I_j(\theta)$
- (B) $I(\theta) = \prod_j I_j(\theta)$
- (C) $I(\theta) = [\sum_j I_j(\theta)][\prod_j I_j(\theta)]$
- (D) $I(\theta) = \int_{\theta} \sum_j I_j(\theta)$

閱讀選擇題第三篇

下面是關於測驗分析之方法，請仔細閱讀後，回答文章後之問題。

有一位國小數學教師，同時教甲乙兩班之學生，學習乘法這個單元。在教完第一個活動後，進行紙筆測驗，此紙筆測驗共有 6 道問題，題目如下：

1. $57 \times 6 = 50 \times (\) + 7 \times (\)$ 2. $78 \times 9 = (\)$ 3. $23 \times 3 = (\)$
 4. $175 \times 7 = (\)$ 5. $14 \times 6 = (\)$ 6. $17 \times 3 = (\)$

又甲、乙兩班之測驗結果如下兩表所示，1 表示答對，0 表示答錯。

試題 \ 甲班學生	A	B	C	D	E	F	G	H	I	J
1	0	0	0	0	1	0	0	0	1	0
2	1	1	0	1	1	0	0	0	1	0
3	1	1	0	1	1	0	1	1	1	0
4	1	0	0	1	1	0	0	0	1	0
5	1	1	0	1	1	0	1	0	1	0
6	1	1	0	1	1	0	1	0	1	0

試題 \ 乙班學生	A	B	C	D	E	F	G	H	I	J
1	0	0	0	0	1	0	0	0	1	0
2	1	0	0	0	1	0	0	1	1	1
3	1	1	0	1	1	0	0	1	1	1
4	0	0	0	1	1	0	1	0	1	0
5	1	1	0	1	1	0	1	0	1	0
6	1	1	0	1	1	0	1	0	1	0

為了分析甲乙兩班之測驗結果，底下提出『難度』、『庫李信度』、『試題順序關聯』之理念，其定義如下：

【定義一】(試題*i*之難度)

假設受試者全體總數為 N ，試題*i*之答對人數為 m ，則試題*i*之難度 P_i 定義為 $\frac{m}{N}$ 。

【定義二】(庫李信度)

假設受試者全體總數為 N ，試題之總數為 K ，第*i*題第*j*位學生之得分為 x_{ij} ，則庫

$$\text{李信度之定義為 } \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K \sum_{j=1}^N (x_{ij} - P_i)^2}{\sum_{j=1}^N \left(\sum_{i=1}^K x_{ij} - \bar{x} \right)^2} \right), \quad \bar{x} = \frac{1}{KN} \sum_{i=1}^K \sum_{j=1}^N x_{ij}。$$

【定義三】(試題順序關聯)

試題*i*到試題*j*有箭頭之充分必要條件為「所有答對試題*j*之學生均答對試題*i*」，記作*i*→*j*或*j*←*i*。若*i*→*j*且*j*→*i*成立時，則稱試題*i*與試題*j*等價。等價在測驗的解釋上，是將兩題所測的概念視為相同。

例如：乙班所有答對第2題的學生均答對第1題，故有2→1或1←2。

請依上文所述回答下列問題：

1. 若根據甲班之測驗結果，則難度最高之題目為何？

- (A) 1
- (B) 2
- (C) 3
- (D) 4

2. 若根據乙班之測驗結果，則難度最低之題目為何？

- (A) 1
- (B) 2
- (C) 3
- (D) 4

3. 根據試題關聯結構之定義與甲班之測驗結果，試題1、試題2、試題4三題間之順序關聯為何？

- (A) 1→4→2
- (B) 2←1→4
- (C) 2→4→1
- (D) 2→1←4

4. 根據試題關聯結構之定義與乙班之測驗結果，試題1、試題2、試題4三題間之順序關聯為何？

- (A) 2→1→4
- (B) 2←1→4
- (C) 2→4→1
- (D) 2→1←4

5. 假設將 $i \rightarrow j$ 解釋為「學會第 j 題的前提是先學會第 i 題」。若根據乙班之測驗結果，則學會第 4 題的前題為學會哪幾題？

- (A) 第 1, 2 題
- (B) 第 2, 3 題
- (C) 第 5, 6 題
- (D) 第 1, 6 題

6. 若根據甲班之測驗結果，則等價之題目為何？

- (A) 第 1, 2 題
- (B) 第 3, 4 題
- (C) 第 5, 6 題
- (D) 第 2, 4 題

7. 根據庫李信度之定義，庫李信度亦可改寫成？

$$(A) \quad \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K P_i(1-P_i)}{\sqrt{S}} \right), \text{其中 } S = \frac{1}{N} \sum_{j=1}^N \left(\sum_{i=1}^K x_{ij} - \bar{x} \right)^2$$

$$(B) \quad \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K P_i(1-P_i)}{S} \right), \text{其中 } S = \frac{1}{N} \sum_{j=1}^N \left(\sum_{i=1}^K x_{ij} - \bar{x} \right)^2$$

$$(C) \quad \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K P_i(P_i-1)}{\sqrt{S}} \right), \text{其中 } S = \frac{1}{N} \sum_{j=1}^N \left(\sum_{i=1}^K x_{ij} - \bar{x} \right)^2$$

$$(D) \quad \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K P_i(P_i-1)}{S} \right), \text{其中 } S = \frac{1}{N} \sum_{j=1}^N \left(\sum_{i=1}^K x_{ij} - \bar{x} \right)^2$$

8. 根據甲班之測驗結果，試求庫李信度？（四捨五入至小數第 2 位）

- (A) 0.23
- (B) 0.34
- (C) 0.8
- (D) 0.92

9. 根據乙班之測驗結果，試求庫李信度？（四捨五入至小數第 2 位）

- (A) 0.23
- (B) 0.34
- (C) 0.8
- (D) 0.92

10. 若數學之邏輯發展程序為『不進位之乘法 → 被乘數的個位進位 → 被乘數的個位與十位均進位 → 被乘數的個位十位百位均進位 → 乘法分配律』，則甲乙兩班中與數學之邏輯發展程序一致之班級為何？

- (A) 甲班是乙班不是
- (B) 乙班是甲班不是
- (C) 甲乙兩班均是
- (D) 甲乙兩班均不是

閱讀選擇題第四篇

我們將在此概述那些與編寫大多數測驗題的技術有關的一般問題，以及那些與編寫特殊類型試題有關的特定問題。最後，我們將描述教師應如何編訂或者使用題庫，以便挑選將編入由他們編製的各次測驗中的試題。

對我們中間的大多數人來言，通過書面文字進行溝通，不是一件輕而易舉的事。試題編寫是殊為困難的一種書面形式，測驗題更是突出。它們不像大多數其他書面形式那樣，能使學生從上下文得到啟發。測驗題是一些比較簡短的陳述；看來，編寫試題侷限於用一兩個句子，比之不受篇幅限制，總是更困難一些。每一道試題都要被大多數學生仔細閱讀，每一個單詞都變得重要起來了，考生不斷地詢問這道試題或者這個單詞真正的意思是什麼。如果試題不是精確地擬定的話，那麼考生就可能把擬題人所不承認的內容硬塞進對試題的理解中去。考試結束後，擬題人就是跟學生再說什麼「你本該知道那並不是我的意圖」，那也沒有什麼用了。教師在編寫試題時可能會出差錯，甚至職業試題編寫者也會出差錯。對於標準化試題的大多數批評，已經把矛頭指向措詞拙劣的多項選擇題。

問題在於，試題編寫跟各項教學目標的製定，同樣都是一個循環過程。不論是編寫客觀性試題，還是陳述性問答題，教師都不應低估編製一道高水準試題的困難，也不應該在編寫的任務面前顯得手足無措，或者感到精神壓力很大。教師需要一開始就盡一切努力，編寫出在技巧上無懈可擊的試題。對於已經編寫好的試題，教師可以通過考試前仔細的複閱、同事們的批評以及學生的反應來加以改進。在考試結束之後，也能夠從試題分析（項目分析）中進一步蒐集有助於改進試題的資料。非正式評鑑也能充實和增補關於某一試題或者某一組試題適合程度的資料。就是說，如果某學生或者若干學生，對於跟某一既定目標有關聯的行為，能夠在課堂上的自發性評量中作出表示，但在書面測驗中卻不能表示的話，教師就應該著手檢查各道書面問題，看一看這些試題之所以困難，是否由試題本身而不是學生或者教學造成的。

在編寫過程中，許多試題將會作廢，而其他的試題能通過事後的認識加以改進。教師不應該因為許多試題編寫拙劣而頓時感到氣餒；這些試題經過評論之後，能夠得以改進。跟所有高水準的寫作一樣，編寫一道測驗題，要求對之推敲、刪改、潤色，甚至有時推倒重來。只要教師記住這個事實：試題編寫並不是臨考前開個夜車就能一次性完成的，而是一個繼續前進的循環過程，那麼一個能用於評量教學並且編寫完善、經實踐證明是可取的題庫，最後終將編訂製成。

因此，本文的重點是闡述改進編寫試題的基本技術方法，關於測量認知與情意分類中的那些不同層次的試題設計技巧，將專門在另一篇文章討論。不過，讀者應該知道，編製一道測量上述任一分類中任一層次的試題，其最為首要的先決條件，就是精心編題。例如，如果一道測量應用能力的試題，由於措詞笨拙而被人誤解的話，或者僅僅按照某一語法線索就能正確地回答該道試題的話，那麼花在設計試題上的時間，就可能白白浪費了。

同樣，我們將要提出各項要求編寫題意明確簡潔、技術上正確的試題建議，對於為常模參照測驗編題，為效標參照測驗編題都一樣適用。對於這兩類測驗中的任何一類來說，題意明確和語法、技術上的考慮同等重要。

最後，在使用題庫的過程中，各項有關編寫試題的一般建議和具體建議，就成為篩選

收入試卷中的各道試題的重要準則。某道試題收入了題庫，並不一定意味著它的編寫有水準。跟我們大家一樣，試題編寫者也可能感染上本文將予概述的各種毛病。教師從題庫挑選各道試題時，需要按照下面扼要敘述的各項一般建議，花些時間加以仔細審閱。

若干初步考慮

各項教學目標與規格明細表

試題編寫的起點，就是對各項教學目標予以仔細考慮。規格明細表（雙向細目表）所概括的各項教學目標目錄，成了在籌劃書面測驗時必須遵循的藍圖。按照教師對於規格明細表中的每一個具體目標小格（細目）的相對總重要性所作出的判斷，可以對每一個小格都賦予一定的價值份量。這種判斷的基礎，可能是對該特定小格所投入的教學時間總量或者所給予的重要性，或者是按照今後教學的觀點所覺察到的教學目標的重要性，或者是所有的小格都被認為同等重要。不考慮對每一個小格加權的各項準則，也可以賦予每一個小格粗略的百分數，並可用這些百分數來估計應為每一個小格編寫的試題數。這種做法有助於確保內容效度，就是說，這個試題反映了各個教學目標的相對重要性。

如果對某個小格賦予一個 20% 的相對份量，而且計畫進行一次包括 50 道試題的測驗的話，那麼為這一個小格將編寫十道左右的試題。按照測驗信度對測驗篇幅作各種考慮，是非常重要的。編寫客觀試題的一般經驗，是編寫者得編寫三倍於所需試題的題目，因為每三個試題中，大概將有兩個會出現這樣或那樣的缺陷（雖然這個比例數隨著實踐將會縮小）。在我們舉的例子中，為了得到十道左右符合要求的試題，教師應計畫編寫三十道試題。一旦各個小格都被賦予了一定的價值份量，而且每一個小格所需要編製的試題近似數目業經確定，那麼就需要先編寫用來引出與每個小格有關聯行為的那些試題。

各類試題的挑選

應該編寫什麼類型的試題來測量上面談及的教學目標呢？按照我們在下一個部分中將見到的情況來看，有兩種基本測驗類型：問答型試題，即學生必須寫下試題答案；選答型試題，即學生從若干個並列的答案中選定正確的答案。我們還將瞭解到，在這兩個範疇下，還有著若干可選用的試題形式。試題形式的選擇，必須服從有待測量的教學目標的性質。應當從一開始就優先選用那種讓學生最直接地展示出合意行為的方法。最後很可能選用較多的間接型試題，不過它們的效度應予確立；換句話說，各人在間接試題上的成績排列的順序，必須跟他們在直接評定方法下所得的成績順序相同。對許多教學目標來說，書面測驗、問答型試題和選答型試題，都是對具體教學目標的直接測量。但是，教師需要考慮到，許多教學目標不可能通過書面測驗直接對之測量——例如，會話技能、外語發音、對實驗室技術的正確使用、顯微鏡的使用或者汽化器的修理。

除了根據是要求學生供給答案還是挑選答案對試題進行分類外，還要根據如何評分的情況，從性質上把各道試題分為陳述性試題，或是客觀性試題。

每一道陳述性試題，須由精通測驗範圍的人士進行審閱。由於沒有一個答案能夠被截然列為純粹正確的答案，因而評卷人員無法對答案進行評分。評分中還可能會攙入對答案質量作出的主觀判斷。事實上，陳述性問答題的主要弱點之一，就是不同的讀者不會對答案的正確性取得一致意見，甚至同一位讀者在不同的時候，也會對各個答案作出不同的評分。

請依上文所述回答下列問題：

1. 請問規格明細表（雙向細目表）的「欄」通常設為什麼？
 - (A) 具體行為
 - (B) 課程內容
 - (C) 能力種類
 - (D) 教學時數
2. 請問規格明細表（雙向細目表）的「列」通常設為什麼？
 - (A) 具體行為
 - (B) 課程內容
 - (C) 能力種類
 - (D) 教學時數
3. 下列何者是最不可能影響規格明細表小格內百分數的因素？
 - (A) 教學目標的重要性
 - (B) 教學時數的總量
 - (C) 命題教師的看法
 - (D) 考生的反應
4. 當規格明細表每一小格的試題近似數目業經確定，接下來該做的事為何？
 - (A) 依照教學內容，定義每一小格的關聯行為
 - (B) 編寫引出每一小格關聯行為的試題
 - (C) 依照具體行為，定義每一小格的關聯行為
 - (D) 挑選適合每一小格關聯行為的試題類型
5. 對於已經編寫好的試題，下列何者最不適合當作教師徵詢改進意見的對象？
 - (A) 家長
 - (B) 學生
 - (C) 同事
 - (D) 自己
6. 下列何者不是試題分析（項目分析）所提供的資料？
 - (A) 難度
 - (B) 鑑別度
 - (C) 效度
 - (D) 選項誘答力

7. 本文提及：教師如欲充實或增補試題的適合程度，可以採用「非正式評鑑」。下列何者不屬於「非正式評鑑」？
- (A) 讓學生反省答錯的原因
 - (B) 讓學生說明或示範其原始的解題方法
 - (C) 再讓學生重考一次
 - (D) 讓學生說明他原先對試題題意的認知
8. 請問下列何者最符合本文所提及的「合意行為」？
- (A) 試題編寫者所欲測量的具體行為
 - (B) 學校教師所教的重點行為
 - (C) 課程綱要所列的教學目標
 - (D) 學生內心所想表達的話
9. 請問下列何者最可能是本文接下來要談的問題？
- (A) 問答型試題
 - (B) 選答型試題
 - (C) 陳述性試題
 - (D) 客觀性試題
10. 本測驗的試題皆為客觀性試題或選答型試題，請問它對能力的測量最偏重於那一種能力？
- (A) 有關事實的知識能力
 - (B) 推論能力
 - (C) 組織能力
 - (D) 閱讀能力