

高中數學之電腦化適性診斷與學習系統

—以矩陣單元為例

俞克斌 施淑娟 許天維

國立臺中教育大學教育測驗統計研究所

摘要

如何偵測學生學習過程盲點，並予以矯正，協助其得以更上層樓，是教育關鍵的一環。故進行高中「矩陣」單元的專家知識結構與對應題庫之建立，並據以為藍本，建置電腦化適性選填題測驗診斷系統。此系統之介面簡單且易於操作，並適用電腦或行動載具進行施測，可減少傳統畫卡式作答與閱卷所造成之人時壓力。並經省題率檢測，足堪確認適性效果，經評估結果顯示：在前測方面，當預測精準度為 96.31% 時，平均施測題數為 16.33 題，大約節省 14.05% 的試題；當預測精準度為 90.25% 時，平均施測題數為 14.80 題，大約節省 22.11% 的試題。後測方面，當預測精準度為 95.79% 時，平均施測題數為 14.10 題，大約節省 25.79% 的試題；當預測精準度為 90.59% 時，平均施測題數為 12.39 題，大約節省 34.79% 的試題。

同時此適性測驗系統，可以診斷高中生矩陣單元學習狀況，鑑別學生個別盲點，輔以針對性補救教學結構，進行適性補救教學。經前測成績校正後的後測成績，電腦化適性測驗組比紙筆測驗組平均高 4.579 分，且達統計上顯著意義 ($P < 0.05$)。各組在前後測成績差異部分，後測的平均成績顯著高於前測逾 20 分 ($P < 0.05$)。此外，經前測成績校正後的後測成績，線上適性補救比傳統團班補救成績高達 10.161 分，統計上達顯著差異。足證「經補救教學後各組前後測成績進步達顯著差異」、「電腦化補救教學優於傳統團班補救教學」，提升高中教學成效可以期待。

關鍵字：電腦化適性測驗、專家知識結構、矩陣、省題率、補救教學。

Adaptive Diagnostic & Learning System based on the High School Math: Matrix Unit

Ko-Pin Yu Shu-Chuan Shih Tian-Wei Sheu

Graduate Institute of Educational Measurement and Statistics,
National Taichung University of Education

Abstract

How to detect the blind spot on learning process of students, and be corrected, to assist them into new heights, was the key point of education. Therefore, we set up the expert knowledge structure and corresponding items bank for each concept of the “Matrix” unit of high school, in order to build up the computerized adaptive optional question test and diagnostic systems. The systems were easy to operate and suitable for computer and mobile use, which could increase manpower flexibility and reduce time-burden. Through the saving item rate test, we found that:

- (1) The pretest average item numbers were 16.33 under the prediction accuracy of 96.31%, at a savings of approximately 14.05% items.
- (2) The average applied to the measured number of items to 14.80 by the prediction accuracy of 90.25%, at a savings of approximately to 22.11% items.
- (3) The post-test average items number were 14.10, when the prediction accuracy of 95.79%, at a savings of approximately 25.79% of the questions.
- (4) Average number of questions were 12.39 when the prediction accuracy of 90.59%, at a savings of approximately 34.79% of the questions.

Further, the adaptive testing system could diagnose high school students' learning situation of “Matrix” unit, and distinguished one's own blind spots, which will be supplemented by targeted remedial teaching. After adjusted by pretest score, the posttest score demonstrated that computerized adaptive testing group is 4.579 higher than paper-based testing group ($P < 0.05$). The groups are characterized by significant differences in pre-posttest which posttest score is 20+ higher than pretest score ($P < 0.05$). Besides, the adjusted posttest score also showed that computerized adaptive test systems is 10.161 significant higher than traditional group remedial teaching. The outcomes given above is sufficient to prove that remedial teaching has significantly helped to enhance learning efficiency, and the effectiveness of computerized adaptive test systems was better than the traditional group remedial teaching style.

Keywords: Computer adaptive testing, Expert knowledge structure, Matrix, Saving item rate, Remedial teaching

壹、緒論

一、電腦化適性測驗

成就測驗原始目的是為了偵測學生之正確能力值，而診斷測驗則旨在偵知學生之學習盲點癥結，俾以對智者彰其精闢、對愚者輔補缺漏，甚爾回饋教學者教學技法之修正。職是之故，顯然能在最短時間、運用最少資源，偵察確認受試者真實能力的測驗工具洵為最佳測驗選擇。欲達成上述目標，則端賴「因材施教」的「適性測驗(adaptive testing)」方得竟其功(劉育隆、曾筱倩、郭伯臣, 2006)。亦即依據個別學生在上一個作答反應的情形，配合合理嚴謹的學理邏輯，來給予下一題適合考生的試題。

生於當今，受惠於電腦科技的興起，將可以利用其強大運算功能，進行電腦化適性測驗(computerized adaptive testing, CAT)。當學生答錯某題時，理論上反映學生真實能力應較該試題對應之能力值為低(除非考量疏忽因素)，無法作答更困難的題目，電腦將會評估呈現適當試題，一則向下偵察學生真實能力，再則使學生不會困窘於超越其能力值之題目，致使學生產生學習挫折，以維繫學生珍貴之學習信心；而對某題答對時，理論上反映學生真實能力有機會較該試題對應之能力值為高(除非考慮猜測因素)，可以嘗試作答更困難的題目，電腦將會評選呈現適當試題，同樣一則向上偵察學生真實能力，再則使學生無須將時間浪費在已經確實可答對的低能力值題目上，這將可以使學生往更高的程度探索發展，激發學生寶貴之學習熱情。準此，電腦化適性測驗確實可以達成節省測驗題數，以致於縮短測驗時間，高效率地偵知受試者之真實能力值(陳世銘、李淑娟、施淑娟、吳任婕, 2008)。

二、以知識結構為主的電腦適性測驗

唯測驗之目的若僅得偵知受試者之能力座標，卻不知道其學習障礙所在，不知道其邏輯盲點所在，則無從「對症下藥」，亦無法達成「適性測驗」的最終極目標，為每一位學員量身打造適才適所的對應補救教材。

以知識結構為基礎的適性測驗(Knowledge Structure based Adaptive Testing, KSAT)與傳統電腦化適性測驗的差別，在於組成學問的「知識結構」(knowledge structure)為基礎所編製的電腦化適性測驗(吳慧珉, 2006)。故其特點就在於先期構築合乎學理的知識結構體系，並據以開發測驗題庫，再循適性選題方式，快速而精確的進行診斷，才能準確鎖定學生的學習困難的概念(節點)，進一步指出造成學習困難的可能原因，如此才能提供教師對受試者設計個人化的補救教學方案。當KSAT的選題策略係透過對「知識結構」脈絡之掌握，則其能偵察受試者知識節點，釐清真正學習盲點之效能，顯然優於單純依難度、能力差異為選題依據的傳統電腦化適性測驗(曾彥鈞, 2007)。

三、以高中生為研究對象

不少學者早已意識到以知識或試題結構為主的電腦適性測驗(KSAT)所能帶來的貢獻，因此歷來多有研究本方法，如：張雅玲(2010)、汪端正(2008)、林婉星(2007)、劉易昌(2007)、曾彥鈞(2006)、趙琬津(2005)，而且獲得許多成果，例如：發現其在診斷學生子技能有無的精準度與節省試題數上皆具有相當高的穩定性與強韌性(黃碧雲, 2005；黃珮璇、王暄博、郭伯臣、劉湘川, 2006)，也解決診斷評量時常面臨到之「電腦適性診斷測驗的編製」、「施測後

的計分及記錄」、「如何根據測驗結果進行適當補救教學」……等棘手問題（郭伯臣，2003；2004；2005）。唯諸位先進所見大抵以國小或國中為對象，少有涉獵高中課程。

其緣由大致有內在與外在兩方面因素，內在者係指高中範圍學理層級遠較國小、國中複雜，單就每一章節知識結構分析便已繁複遠逾國中小，遑論據以編製符合實驗理想之試題；外在者則係指高中生自主性高，配合實驗意願低，同時歷經國小、國中之分化，致使學員程度歧異大，若欲完整收集答題反應類型，進而全盤探究各成員素質典型，實為浩瀚工程，凡此均使企圖針對高中階段作學術研究者望之卻步。但相對而言，就因為高中學生學習成效評估的系統性、規模性探討付之闕如，使得銜接而上的大學高等教育必須在摸索中調整，教學雙方迭感困窘。此等疏遺，亟待補足（俞克斌、施淑娟、郭伯臣、鄭俊彥，2012）。

四、高中課程矩陣單元

「矩陣」在高等教育與實用領域的地位舉足輕重。矩陣是高等代數學中的常見工具，也常見於統計分析等應用數學學科中。在物理學中，矩陣於電路學、力學、光學和量子物理中都有應用；計算機科學中，三維動畫製作也需要用到矩陣。矩陣的運算是數值分析領域的重要問題。將矩陣分解為簡單矩陣的組合可以在理論和實際應用上簡化矩陣的運算。即便在高中數學教學中，國內教育專家亦從未輕忽，自民國60年代迄今，均在高三課程中佔有一席之地，（中國大陸自2004年起亦順應潮流，第一次將「矩陣」編入高中課程中），甚至根據中華民國97年1月24日教育部發布的普通高級中學「數學」課程綱要規範（簡稱99課綱），將「矩陣」章節提前至高二學習，「矩陣」單元之為重，可見一斑。由是觀之，標舉「矩陣」單元作為研究標的，尙有價值。

五、適性化補救教學

現今數學教育實況與教育理想相去甚遠，而且陷入一種惡性循環：先是學生紙筆運算能力低落、文字理解能力下降（幸不論其源頭是否為學習歷程之更前端，過度強調減輕學生作業壓力，過度強調圖像學習趣味性），以致單位時間練習量下降，嫻熟度下降，不得已仰仗校外補習班補強，更減低學生每週可自由運用時間，以致練習量再度下滑，由是每況愈下，終至伊於胡底。

而光是抱怨不足以解決問題，打破惡性循環之道容或千頭萬緒，唯最有效率者無非「適性補救」一途，蓋短時間內企圖提升學生計算速度、拔擢學生文字理解均有困難，唯獨偵知學生之「知識結構」並配合「適性補救」，利用其針對個別學習障礙差異對應性予以補強，卻能夠節省時間耗損之特質，方能在裨補缺漏之餘，不再增加學生負擔，尙為最方便法門。

準此以觀，若能妥善規劃建置一個「以知識結構為主的電腦適性測驗的高中數學診斷系統」連同「補救教學系統」，並在實行面獲得驗證，對於釐清高中學生學習矩陣單元之盲點，提升高中教師教學果效，均可預期多有裨益。

根據上開背景，本研究的企圖探尋之研究目的如下：

- （一）建立高中「矩陣」單元的專家知識結構與對應題庫，俾以建置電腦化適性測驗系統。
- （二）運用電腦化適性測驗系統，診斷高中生矩陣單元學習狀況。並以全測資料進行模擬實驗，檢測省題率以確認適性效果。
- （三）根據適性測驗所得，鑑別學生個別盲點，進行適性補救教學，並探討電腦化適性測驗系統補救教學效果。

貳、文獻探討

一、知識結構

張新仁(1993)定義知識結構是在腦中的一個有組織的層級架構，而此架構是由學習者將所學得的知識所組成。涂金堂、林佳蓉(2000)進一步認為此結構是學習者將所接收到的外界訊息，在腦中消化、吸收及重新組織而成的。而知識結構的功能、影響，據 Bower (1972) 及 Shavelson (1972) 的研究，可以透過建構、修正和重組知識結構，來影響學習和認知的表現。雖然不同的學者，對於人類知識表徵的方式有不同的解釋，但大致均有一個共識，即在特定的知識領域內，概念與概念之間，都具有某種交互連結的關係，因而形成一種組織上的特性，此種概念與概念間的組織特性，即可稱之為「知識結構」。

在知識結構中，上層的結構是屬於比較困難的，稱為「上位概念」，下層的結構則為各概念的「下位概念」(黃碧雲, 2005; 曾彥鈞、劉育隆、郭伯臣, 2006)。

從相同知識領域的不同知識結構的不同組織特性中，可看出不同專熟程度，例如：Larkin, McDermott, Simon & Simon (1980) 早就在比較物理領域的專家與新進人員在解決動力學問題上的行為表現，發現了專家與新手在知識組織方面和解題策略方面確實有所差異。再根據 Chi, Glaser & Rees (1982) 的研究，發現專家在某特定領域的知識結構，可能比一個初學者更為一貫、更為簡捷，甚至更為抽象。Shavelson & Stanon (1975) 觀察學生與教授在物理學名詞類聚結構上的異同，發現隨著練習的逐漸增加，學生的名詞類聚結構是可以逐漸接近老師的名詞類聚結構，Thro (1978) 亦早有類似的發現。

既然知識結構的特性可以反應出不同的專熟程度或學習狀態，因此如何瞭解或評量個體的知識結構，便成為值得深入探討的課題。

(一) 專家知識結構

就教學而言，專家知識結構是由學者專家根據學理及經驗，分析教材內容及教學目標，找出單元內重要的學習概念，再根據學生的學習歷程、學習順序及概念發展的關係所整理建立而成。本於 Hall (1988)、Rewey, Dansereau, Skaggs, Hall & Pitre (1989)、Skaggs (1988) 等的看法，透過專家產生的知識結構，可以讓學生獲得概念鷹架 (concept scaffold)。甚至知識結構專家化的程度，可以有效預測學生的學習表現 (江淑卿、郭生玉, 1997)。

根據吳慧珉 (2006) 的研究，首先遴聘數學教育學者及數名現職教師組成專家會議，根據教學學理以及專家成員的教學經驗，共同分析各個目標單元施測範圍內的知識概念，配合學生的學習歷程、概念發展順序及其概念間上下位關係的合理性，整理建立一組完整且具有階層性概念的數學知識結構關係，即為專家知識結構。專家知識結構並非唯一且絕對，不同的成員可能討論出不同的專家知識結構。

吳慧珉列出設計專家知識結構時之原則：

1. 每一節點為單一概念。

2. 了解上位節點則推論受試者了解其所有下位節點。
3. 同層節點難度可能不一。
4. 注意數字大小對難度所造成的影響。
5. 注意文字題及非文字題的不同。
6. 當無法明確界定節點間順序性時，宜定義為無次序性。

最後依據吳慧珉（2006）所編製的專家知識結構檢核表，逐一檢視該專家知識結構是否符合規定，如表 1 所示。

表 1 專家知識結構檢核表

項目	檢核內容	專家判斷
1	每一個節點是否都是單一概念？	是 ☐ 否 ☐
2	上位節點和下位節點的次序性是否明確合理？	是 ☐ 否 ☐
3	節點的設定是否有考慮到數字位數大小的問題？	是 ☐ 否 ☐
4	節點的設定是否有考慮到文字題和非文字題的不同？	是 ☐ 否 ☐
5	節點的敘述是否清楚明確？	是 ☐ 否 ☐
6	節點的設定是否有符合課程資料的教學目標？	是 ☐ 否 ☐

圖 1 表示一個高中有理數與無理數課程的專家知識結構。專家知識結構亦可作為編製試題的標準，經由測驗了解學生的學習情形。

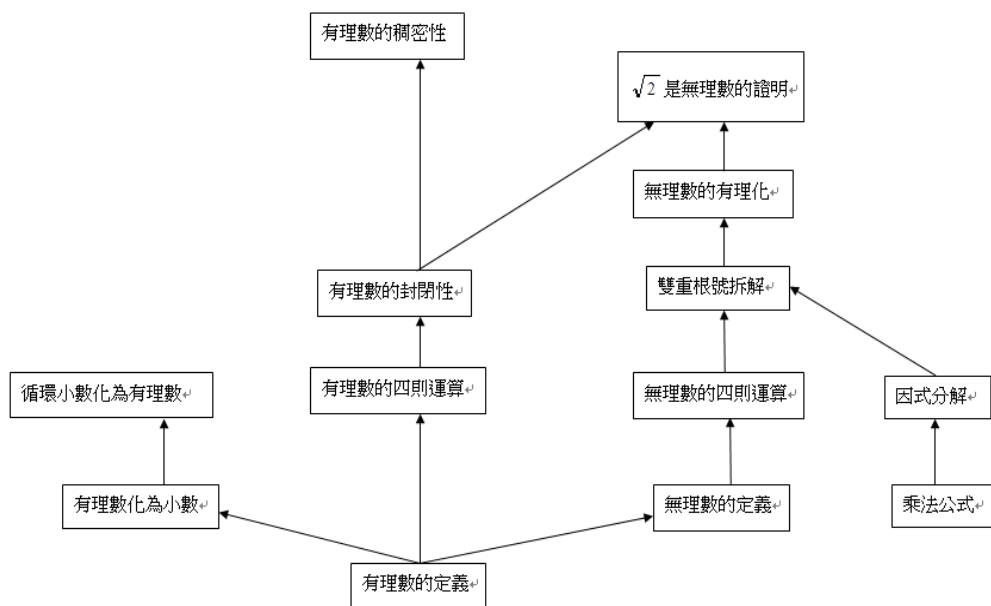


圖 1 高中有理數與無理數課程的專家知識結構

（二）學生試題結構

當我們根據專家知識結構中之知識節點編製出測驗試題，並在學生測驗後，依照其作答反應情況，運用試題結構演算法計算，再加上概念發展順序及概念之

間上、下位關係，將會組成一個學生試題結構。而學生知識結構的建立則是根據學生試題結構進一步分析而成，並且往往需要靠專家依手動方式判斷（黃碧雲，2005；劉安哲，2007）。

學生試題結構和專家知識結構不一定相同，這是因為教學者根據較完整的學理背景與實務印證所建築的體系，與學生經由教學者提供的有限素材，以及認知消化後的引領，所理解領略刻畫的體系，不盡然相同，就像是親身旅遊者與聆聽旅遊記者，感受不盡然相同所致。

圖 2 表示一個高中有理數與無理數課程的學生試題結構。

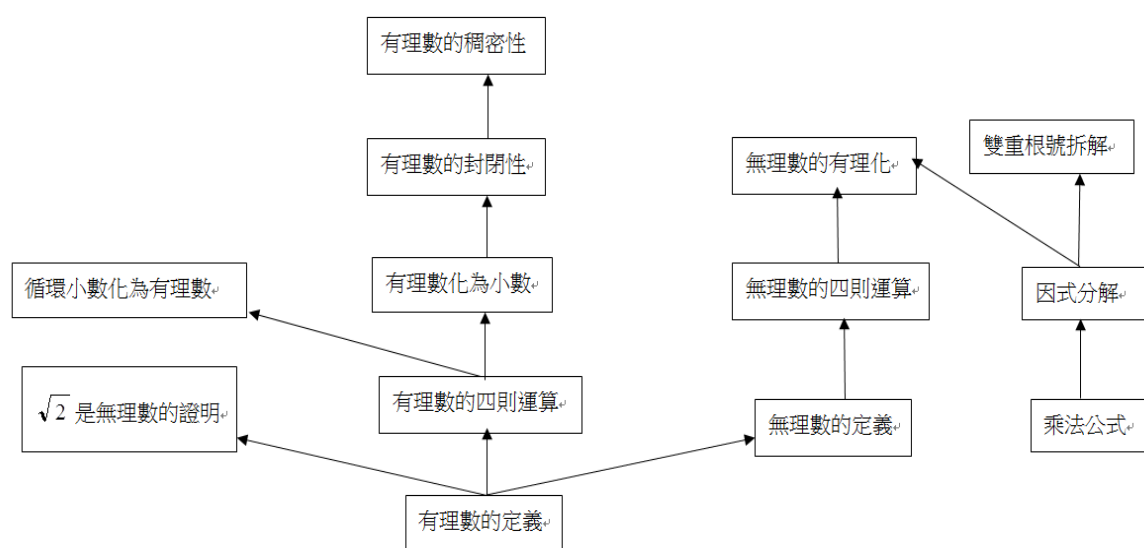


圖 2 高中有理數與無理數課程的學生試題結構

Appleby, Samuel & Treasure-Jones (1997) 曾經一針見血地點出由於數學知識具結構性及階層性，學生進行電腦適性診斷測驗所需作答的試題遠較一般紙筆診斷測驗少。

二、電腦化適性測驗

電腦化適性測驗的實施程序一般包含開始、繼續、以及結束三大步驟(Wainer, 1990)，本身就是個綜合策略考量：起始題如何決定，能力值估計依據，向上向下選題準則，以及終止條件之設定，均需於施測前即先確定依憑法則。

根據郭伯臣、劉湘川、許天維在 2004 年的研究，電腦化適性測驗依理論基礎大致可分為二大類：一類是以試題反應理論 (Item Response Theory, IRT) 為基礎 (Wainer, 2000)；另一類則是以知識或試題結構為基礎 (Appleby, Samuels & Treasure-Jones, 1997; Brown & Burton, 1978; Chang, Liu & Chen, 1998; VanLehn, 1988; Wenger, 1987)。

(一) 以試題反應理論為基礎的電腦化適性測驗

以試題反應理論為基礎的電腦化適性測驗，是將施測後受試者成績視為一「能力值」(ability) 或「量尺分數」(scale score)，較適合用於以教育資源分配

作目標之情境，例如：基本學力測驗、大學入學測驗等。其缺憾則在於僅得診斷出受試者之能力座標，卻無從偵知其錯誤類型（郭伯臣、謝友振、張峻豪、蔡坤穎，2005；劉育隆、曾筱倩、郭伯臣，2006），即無法「對症下藥」，遑論為每一位學員量身打造適才適所的對應補救教材。

（二）以知識及試題結構為基礎的電腦化適性測驗

早在 1998 年，Chang, Liu & Chen 就曾設計一個診斷迷思概念之測驗系統。該研究的重點在於每題的選項除正確選項之外，其餘選項皆被設計成與 M1、M2、M3、M4 等四種迷思概念之一有關，如圖 3 所示。因此，若學生作答第 1 題時，選擇 1a 即可判斷學生可能擁有某一迷思概念且與第 2 題有關，故必須再進行第 2 題的測驗來繼續作答。此時若學生面對的選項為 2a、2b、2c，且這三者分別代表不同的迷思概念 M1、M2、M3，因此即可判斷出學生所擁有的迷思概念為何。若學生選擇的是 1b，則可對應到迷思概念 M4，唯有選擇到 1c 時才能確定學生沒有 M1 至 M4 這四種迷思概念（曾彥鈞，2007）。

此方法優點為：1.根據受試者作答之選項來決定下一題試題，符合適性化之旨趣，可達成節省施測題數，節省測驗時間目標。2.由受試者作答之選項來推論迷思概念，如果題目間脈絡關聯確實，節點間因果關係確切，將可精確診斷出每一位受試者之迷思概念，達到個別化之需求。

而其缺點則為：1.為求精確診斷出相對應之迷思概念，每條適性測驗路徑皆須精心設計以避免重複。2.出題必須考慮選項及迷思概念間關係，並兼具試題上下位難度評估，著實費時費力，困難度高。3.出題必須嚴謹排除可猜測性，否則不易達成精確診斷。

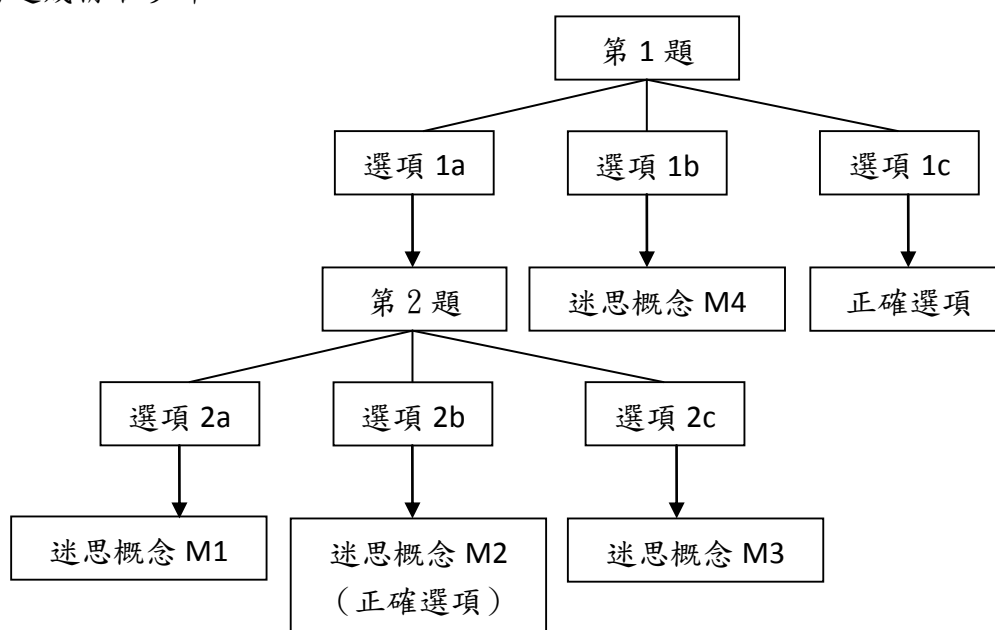


圖 3 診斷迷思概念之測驗系統（Chang et al, 1998）

而郭伯臣(2003;2004;2005)在以「國小數學科電腦化適性診斷測驗(I)(II)(III)」為題的國科會專題研究中，就是秉承 Chang 團隊的研究理路，主張選擇以順序

理論(OT)為基礎所建立試題和試題之間的上下位關係,確認節點間因果關聯,進而組成切中學理之試題結構乃至知識結構,如圖4所示。

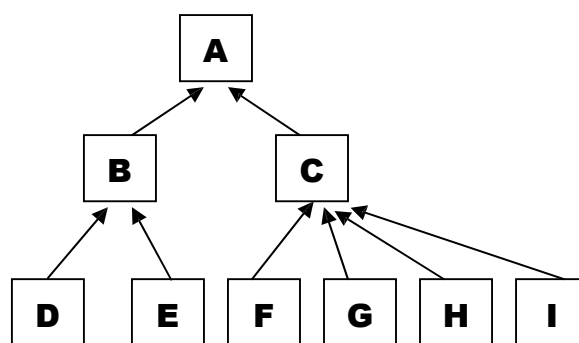


圖4 知識結構

圖中上層的A表示B跟C的上位試題,則相對地B跟C就是A的下位試題。郭伯臣以國小數學為例:A試題是四則運算的題目,而B跟C是加減乘除的題目。當學生能夠答對上位的四則運算題目,則能引喻下位的加減乘除的概念確已精熟。

相較於Chang團隊的成績,此研究所推衍的進步在於加入了順序理論(OT)規範,使適性目的更易達成。因此郭伯臣的研究顯示,傳統的紙筆測驗不論受試者能力值有否軒輊,學習障礙是否差異,受試者均必須作答完整的A~I共9題;然則若藉著已循順序理論規範的電腦化適性測驗系統施測,只需先讓受試者作答對應上位概念的試題A,設若受試者答對則表示其下位概念理應均屬精熟,亦即無須再耗時作答,此時節省題數即B、C、D、E、F、G、H、I等8題,反之若試題A作錯時則接續作答B、C兩題,如果發現試題B答對而試題C答錯時,電腦將會判定受試者對試題D和試題E的概念是已經學會了將可不必再作答,而省下這二題的作答時間,但受試者必需再繼續作答試題F、G、H、I,進一步釐清知識概念之盲點。

此研究可謂落實Chang團隊理論精髓暨適性選題之宏旨:既可達偵察受試者學習障礙所在,診斷出學生不同錯誤概念之旨趣,亦可收節省施測題數,節省測驗時間,達成適性化的目標之果效。

本於嚴謹順序理論佐證、基於明確知識結構規範所建置的KSAT系統,適足以精準地緝捕學生學習過程不自知之盲點,協助教學者進一步提供適當之補救,洵為節省人力物力乃至時間的最理想測驗工具。

三、補救教學

當我們藉著各種診斷工具或理論已偵測出受試者之學習弱點或知識盲點,則補救教學即是刻不容緩的後續工作,適當的補救教學可以有效率地幫助低成就學生解決學習上的問題,以提高學生學習的成果。

補救教學的實施理當依據評量測驗後之分析結果來進行,並應輔以專家對受試者個人的綜合評量,進一步理解學生既有的知識與經驗,藉此可從中找出學生

易發生的錯誤類型，發現其學習上的問題並加以輔導修正。王惠美（2002）也提到許多學習評鑑系統在評估學生學習成果時，往往只是計算答題結果，顯示學習者答對或答錯的題數，或僅止於診斷學習者的錯誤型態，並無法進一步地分析每一種概念迷失背後的原因。基於上述補救教學原則，適性補救教學的模式便應運而生。

根據林立敏、白曉珊、郭伯臣、劉育隆（2007）提出適性補救教學可分為二種：一是以電腦為基礎之適性補救教學（computer-based adaptive remedial instruction, CBRI），另一則是以教師為基礎之適性補救教學（teacher-based adaptive remedial instruction, TBRI）。以下就針對二種類型的適性補救模式加以說明。

（一）以電腦為基礎之線上補救教學

既然透過適性診斷測驗可精確的分析出不同能力的受試者所呈現的不同學習困難和迷思概念，電腦化適性補救教學便可達成補救教學中適性化與個別化的目標。

首先以知識結構為基礎，依據專家知識結構圖中的概念節點進行設計，製作符合概念節點的補救教學教材。更因著各種電腦行動化載具的開發，進而讓受試者有機會能進行即時線上補救教學，唯尚待克服之缺點是學生在進行補救時所衍生的後繼問題可能無法立即解決，且教學現場必須準備足夠的軟、硬體設備，對現行教學場域來說亦有實質困難。

（二）以教師為基礎之個別補救教學

以教師為基礎之適性補救教學，是以人力進行補救教學，教師必須在學生進行測驗後針對學生的錯誤概念，配合學生需求，以教材和學生練習本等工具進行補救教學，此種模式教師能依學生個別差異給予立即指導。

洪素敏（2004）研究指出老師適當的介入，給予線索、協助探索，是有正面的教育意義的。其一，可以減少學生的嘗試錯誤率；其二，減少浪費太多的時間；其三，讓學生能朝正確的歷程邁進。因此，教師在補救教學中的角色，結合具體與半具體的操作學習於整個補救教學活動中，給予學生多方面的感覺經驗，從瞭解、互動與溝通到學習的引導，是電腦所無法達成的。就補救教學的本質來說，可以看做是教師教學活動的一部分，而這個教學措施是適應學生個別差異的具體實踐，也可以看做是「適性化教育」的落實（施淑娟、吳玫君、許天維、陳淑勤，2008）。

上述兩種補救方式就學術原理而言，似乎各有短長，寧無軒輊。但本研究企圖透過實驗比對方式，探究對實驗樣本而言兩種學理之落實是否有明顯之差異。差堪可以反應對目今教育場域所訓練出之學子特質，兩種學理究竟孰優孰劣。

四、高中課程矩陣單元

中華民國 97 年 1 月 24 日教育部發布的普通高級中學「數學」課程綱要（簡稱為 99 課綱）一開始就明白揭橥：矩陣是線性代數、離散數學、多變量微積分、多變量統計分析的基本工具，並強調自然組與社會組的學生都需要。

99 課綱規範矩陣的教學應涵蓋一般矩陣的加法、純量乘法與乘法。並可涵蓋一般線性方程組，介紹高斯消去法，並配合實用的例子。其次介紹矩陣的應用則介紹轉移矩陣與二階反方陣，此部分的學習應與線性組合相連結。最後介紹二階方陣的線性變換意涵。

具體臚列 99 課綱中高中矩陣領域的教學綱目：

- (一) 矩陣的定義（含轉置矩陣、對稱反對稱矩陣）。
- (二) 矩陣的加減（含純量乘法）。
- (三) 矩陣的乘法（不談特徵方程式及 Caley-Hamilton 定理）。
- (四) 馬可夫矩陣（即轉移矩陣）。
- (五) 反矩陣（含二階反方陣、三階反方陣，並與解聯立方程組作連結）。
- (六) 高斯消去法（含矩陣的列運算、矩陣的三角化）。
- (七) 平面上二階方陣的線性變換（含伸縮、旋轉、鏡射、推移，及線性變換的面積比）。

五、選填題題型

為精準反應未來學生面對大學學力測驗之形式，本研究使用仿大學學測之選填題題型。

選填題是為了讓學生作答反應可以使用電腦讀卡，將答案分成數格，每一個格子可以填「±」、「—」及數字「0」、「1」、「2」、...、「9」，如圖二中第 1 題，答案為兩位數字，或是一個正、負符號加上一位數字，故有兩個格子「①②」，若答案為「—3」，則在「①」的位置畫卡畫「—」，「②」的位置畫卡畫「3」；若答案為「28」，則在「①」的位置畫卡畫「2」，「②」的位置畫卡畫「8」。選填題之作答說明如圖 5（徐發興，2011）所示。

例 1. 某題的答案格式是 $\frac{\textcircled{20}\textcircled{21}}{50}$ ，而答案是 $\frac{-7}{50}$ 時，則正確畫記方法是：

分別在答案卡的第 20 列的 與第 21 列的 畫記如下：

20	1	2	3	4	5	6	7	8	9	0	-	±
21	1	2	3	4	5	6	7	8	9	0	-	±

圖 5 選填題作答說明

既瞻乎先進肇宏已如上述，則本研究預擬本砥 99 課綱，選擇對承啟高中與大學教學特具意義之「矩陣」單元為研究標的，並依循順序理論（OT），建置高中矩陣單元的專家知識結構與對應題庫。

進而以知識結構為主的電腦適性測驗（KSAT）為背景架構，以選填題形式建置電腦化適性選填題測驗診斷系統，俾以偵測學生答題省題率及學習盲點，並

與傳統紙筆測驗評量，鑑證此系統於高中課程領域之廣泛適用性。

一俟偵測得學生學習盲點，進而輔以針對性補救教學結構，建置對應線上補救教學課程系統，進行適性補救教學，再與傳統團班補救教學作實驗比較，企圖探究經補救教學後各組前後測成績進步是否達顯著差異，期待藉此提升高中教學成效。

叁、研究設計與實施

一、研究流程

在確定研究主題後，針對 99 課綱所臚列之高中矩陣領域歸納出該範圍之知識節點，並經專家會議審定。再進行對應試題編製，亦經專家修審以及 60 餘位同學之預試，偵測試題適用性，再由專家檢核確定試題入庫。一俟題庫確定，即據以編製對應補救教材，並預先錄製線上補救課程。同時設計電腦化適性診斷系統介面。最後篩選受試樣本，依據「前測」、「補救」、「後測」流程進行施測。正式施測結束後進行分析，最後撰寫研究結果。相關研究流程及架構，如圖 6 所示。

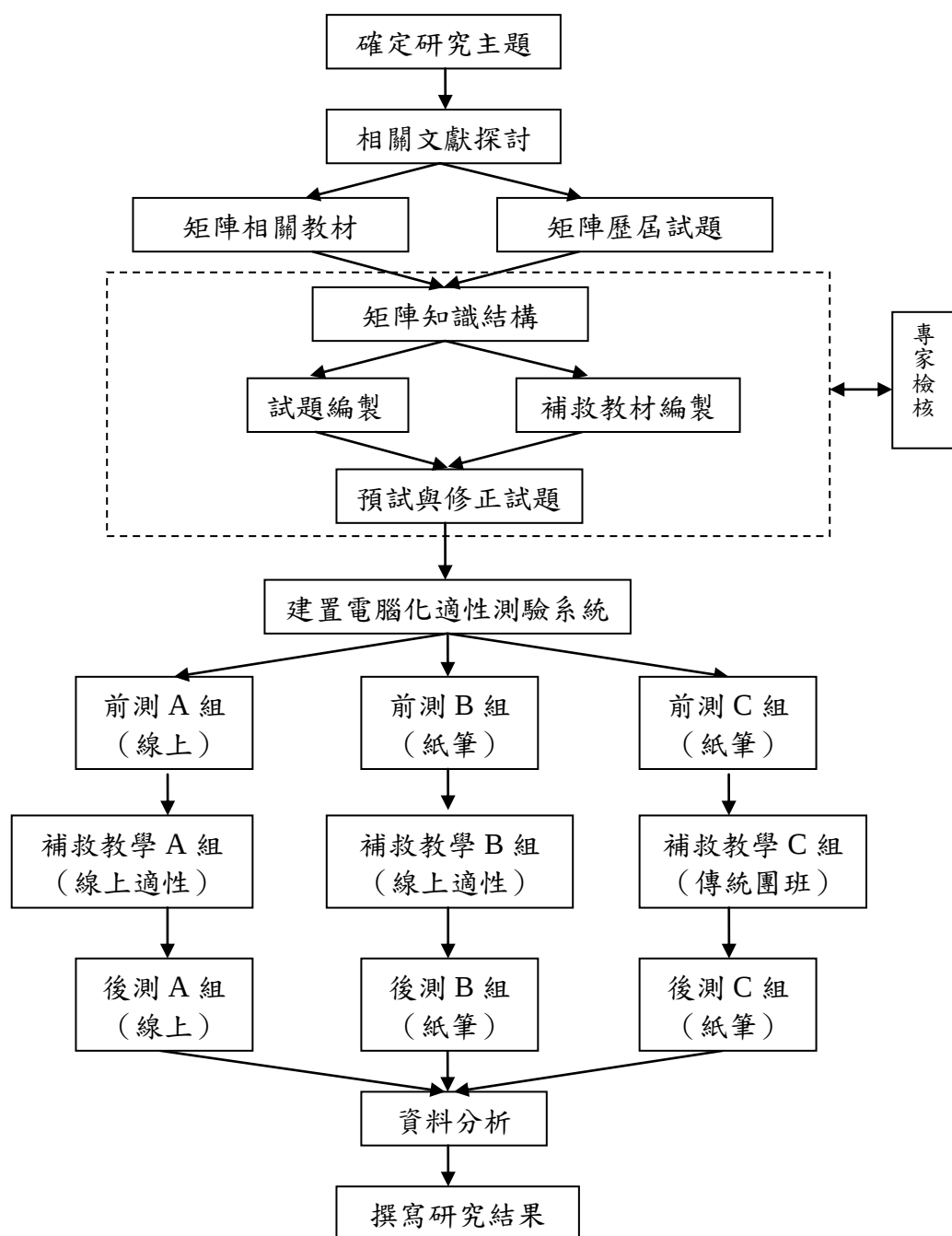


圖 6 研究流程與架構

二、開發對應專家知識結構之線上適性測驗系統

(一) 專家知識結構建立

根據課綱及各版本教科書「矩陣」章節中的說明與範例，並經專家會議討論，探討課綱規範，楹列共同必授課題及所對應知識節點，汰除特定版本教材課外衍生補充，獲得 20 個共同主題，藉以歸納出「專家知識結構圖」，如圖 7 所示。

(二) 編製對應知識結構之試題

根據 Prosser (1974) 的建議，在建構題庫時，每個概念至少要包含 10 個試題，每一單元課程內容至少要包含 50 題。而 Reckase (1981) 則主張一百至二百

個難度均勻分佈，且具有合理鑑別度的試題，便可適用在電腦化適性測驗裡（引自余民寧 2009）。據此標準，針對上述知識節點，每一主題建構包含 10 題以上確實可用之對應試題。同時篩選 62 位受試者進行預試，探討各測驗試題的適切性。此些受試成員與正式施測成員組成成分相當，學生就讀於高雄縣市各公私立高中，素質均勻涵蓋各能力值。預試結果交專家會議，過濾不良試題，補入適當取代題，並再次施測，直至全體試題均合適本研究為止。據此要領所建立之電腦化測驗試題 Cronbach's Alpha 值為 0.807，信度良好。

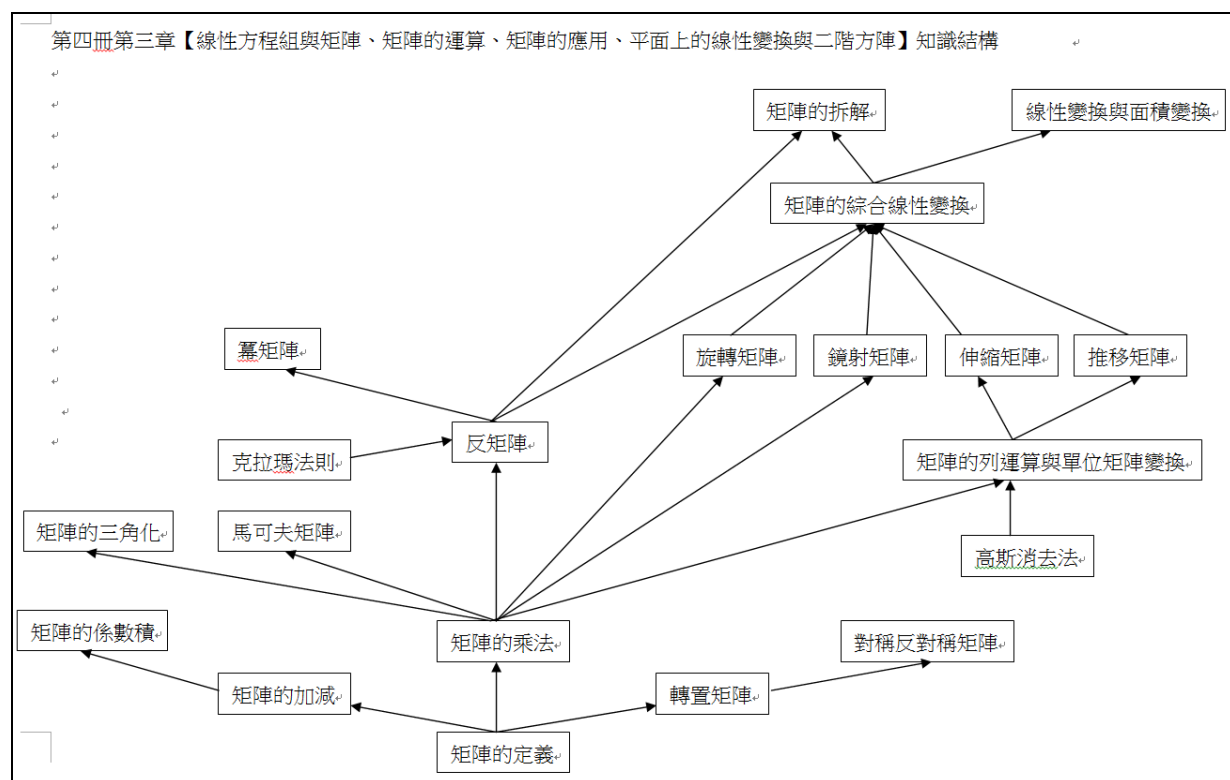


圖 7 專家知識結構

(三) 線上適性測驗系統開發

當題庫建構完成，則據以編輯對應試題。為使補救教學效果精確呈現，前後測試題採複本測驗方式編製。考量學生適應大學學力測驗之形式，系統使用仿大學學測之「選填題」題型，其介面使用 PHP (Personal Home Page) 語言撰寫，畫面如圖 8 所示。

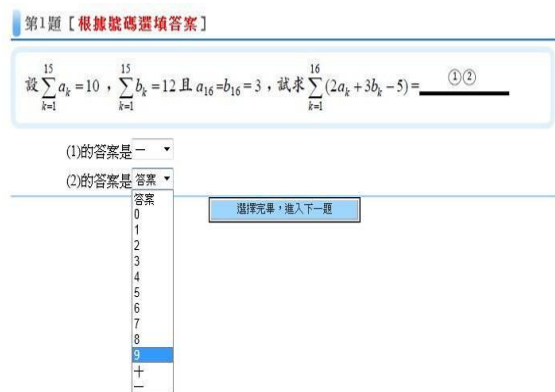


圖 8 線上適性診斷測驗系統
(測驗題頁 1)



圖 9 線上適性診斷測驗系統
(首頁)

學生閱讀題目後利用紙筆計算，並依序選出正確的答案，以圖 8 問題為例，若學生計算的答案為「-9」，則在「(1)的答案是」的位置選擇「-」，並在「(2)的答案是」的位置選擇「9」，確認無誤後，再點選「選擇完畢，進入下一題」。

本研究藉由國立臺中教育大學教育測驗統計所研究團隊偕同毅夫文教機構所開發的「適性測驗系統」，建置研究之實驗過程電腦線上適性測驗系統。如圖 9。

系統依使用者的角色有不同的權限。教師為測驗的指導者，亦可為系統之管理者，除了在系統中負責帳號管理、修改個人資料外，更可以藉由「建立單元知識結構」、「建立試題庫」進行題庫編修之預備及試卷存取之管控，俾能進行施測。圖 10、圖 11、圖 12 分別展示管理者介面、建立單元知識結構介面、建立試題庫介面。



圖 10 線上適性診斷測驗系統
(管理者介面)



圖 11 線上適性診斷測驗系統
(建立單元知識結構介面)

教師尚可使用參加測驗、成果查詢（班級或個人）、系統公告、使用說明等功能。如圖13。



圖 12 線上適性診斷測驗系統
(建立試題庫介面)

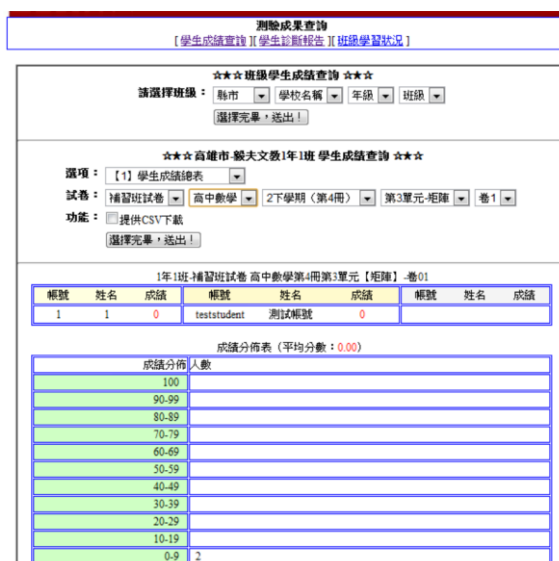


圖 13 線上適性診斷測驗系統
(成果查詢介面)

至於受試者透過網路，輸入個人的帳號、密碼登入系統後，再選取施測單元開始施測，施測完畢後受試者有權限可點選成果查詢。

以下為系統各步驟說明及執行畫面：

1. 登入系統：受試者會有專屬帳號、密碼以登入系統施測，其入口介面與管理者相同。如圖9所示。
2. 點選施測單元：登入系統並點選「參加測驗」鍵，則每位受試者會進入到選取測驗單元介面，此時系統畫面會出現測驗的注意事項及使用說明等提示語，如圖14所示。



圖 14 線上適性診斷測驗系統
(說明頁)

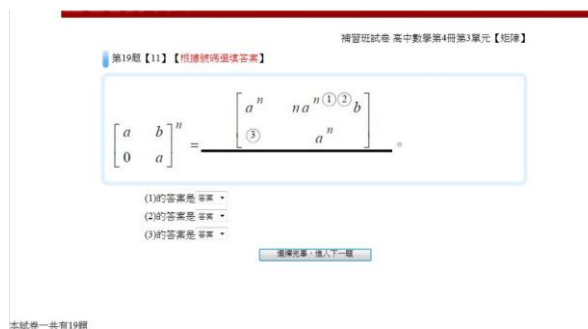


圖 15 線上適性診斷測驗系統
(測驗題頁)

3. 參加測驗：本研究及後繼研究預擬建構全面涵蓋高中三年六冊課程之完整題庫，故介面中已有多份不同單元試卷可供點選，由受試者點選合適試卷受測。如圖15所示。
4. 成果查詢：經系統施測完畢後，學生及教師可點選「成果查詢」功能，系統會透過學習診斷報告書，列出受試者歷來施測單元及學生的答題反應，以作為學生的自我了解，和教師尋求教學補救與改善的參考，如圖16及圖17。

學習診斷報告書

親愛的【測試帳號4】小朋友，你對於【矩陣】單元的學習表現得可圈可點，值得嘉許，相信持之以恆，你也可以成為一位了不起的數學家囉！

基本資料

就讀班級	一年一班	測驗學號	4
學生姓名	測試帳號4	學生性別	男
縣市地點	台中市	就讀學校	台中教育大學

基本資料

補習班試卷 高中數學第4冊第3單元【矩陣】

次數	測驗日期	時間	成績	題數
1	2012-06-15, 23:39	226秒	100	19題/19題

圖 16 線上測驗學習診斷報告
(測驗紀錄頁 1)

學習概況分析圖		
說明： 1.●表示通過該概念；×表示未通過該概念。 2.線上學習教材容量較大，網路傳輸過程中請耐心等待。		
概念列表	診斷結果	線上學習
【重點1】矩陣的定義	●	教材1
【重點2】轉置矩陣	●	教材2
【重點3】對稱反對稱矩陣	●	教材3
【重點4】矩陣的加法	●	教材4
【重點5】矩陣的係數積	●	教材5
【重點6】矩陣的乘法	●	教材6
【重點7】馬可夫矩陣	●	教材7
【重點8】高斯消去法	●	教材8
【重點9】矩陣的三角化	●	教材9
【重點10】矩陣的列運算與單位矩陣變換	●	教材10
【重點11】反矩陣	●	教材11
【重點12】零矩陣	●	教材12
【重點13】旋轉矩陣旋轉矩陣	●	教材13
【重點14】複射矩陣	●	教材14
【重點15】伸縮矩陣	●	教材15
【重點16】推移矩陣	●	教材16
【重點17】矩陣的綜合線性變換	●	教材17
【重點18】矩陣的分解	●	教材18
【重點19】線性變換與面積變換	●	教材19

圖 17 線上測驗學習診斷報告
(測驗紀錄頁 2)

目前讀者可以使用 <http://210.240.193.249/ifkast>，配合帳號:1/密碼:1，實際進行測試，具體感受施測流程。

三、開發線上適性補救教學系統

為偵測「線上補救教學」與「團班補救教學」所導引出之學習成果是否有顯著性差異，事先針對「矩陣」單元，除編製對應知識節點之書面補救教材外，並分別採用「一般錄影」模式錄製「矩陣觀念講解課程」，採用「Camtasia Studio」模式錄製「矩陣練習講解課程」，提供 A 班（紙筆測驗前測）、B 班（線上適性前測）作為學生前測答錯時之適性補救教學。並與 C 班之「傳統團班補救教學」做出對比。

Camtasia Studio 是個螢幕影像攝錄、轉檔、編輯的影片工具，使用者在螢幕上的所有動作或是螢幕上出現的畫面、動畫都可以藉此錄成影片，也能包含聲音或音訊，錄製好之後的片子更可以在軟體中加入文字、指標、或是圖片來加強影片的說明效果。唯其與一般錄影最大差異在於並無教學者人像之呈現，純粹凸顯解題步驟。

為顧及學生自動適性補救效果，每則線上補救課程均以 5~10 分鐘講解時間為限。



圖 18 線上補救教學（適性補救）
一般錄影



圖 19 線上補救教學（適性補救）一
般錄影

設 M 為一 3×2 矩陣，而且 $M \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 2 & 0 \\ 1 & 0 \end{bmatrix}$

， $M \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 2 \\ 0 & 3 \end{bmatrix}$ ，則 $M \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \underline{\hspace{2cm}}$ 。【90 日自】

答： $\begin{bmatrix} 1 & 3 \\ 2 & 2 \\ 3 & 1 \end{bmatrix}$

解：設 $M = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$ ，則 $M \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 \\ c & 0 \\ e & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 2 & 0 \\ 1 & 0 \end{bmatrix}$ ， $M \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & b \\ 0 & d \\ 0 & f \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 2 \\ 0 & 3 \end{bmatrix}$

圖 20 線上補救教學（適性補救）
Camtasia Studio

設 $A = \begin{bmatrix} 2 & -1 & 3 \\ 2 & 0 & 4 \end{bmatrix}$ ， $B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 8 & 4 \end{bmatrix}$ ，求 AB 。

答： $\begin{bmatrix} 26 & 15 \\ 34 & 20 \end{bmatrix}$

解： $AB = \begin{bmatrix} 2 & -1 & 3 \\ 2 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 8 & 4 \end{bmatrix} = \begin{bmatrix} 26 & 15 \\ 34 & 20 \end{bmatrix}$

圖 21 線上補救教學（適性補救）
Camtasia Studio

四、進行實驗

（一）實驗分組

實驗選擇高雄某補習班高二數學班全體 544 位學生進行施測（已刪除無效樣本），由於該補習班學生成員包括高雄縣市所有公私立高中，故其素質可視為均勻涵蓋各能力值。

為比較不同測驗形式與不同補救教學類型之搭配，所導引出之學習果效有否差異，本研究將學生分成三組，三組之測驗形式與補救教學類型分配如表 2。由於學生能力素質均勻，並配合補習班教學環境，採部落抽樣法分配組別取，至於第一、二組人數則因施測場地電腦設備數量限制而得出。並為確認實驗結論，使補救教學效果精確呈現，前後測試題採複本測驗方式編製。

表 2 三組測驗形式與補救教學類型分配

組別	受試者人數	前測方式	補救教學	後測方式
A 組	125 人	線上適性測驗	線上適性補救教材	線上適性測驗
B 組	121 人	紙筆測驗	線上適性補救教材	紙筆測驗
C 組	298 人	紙筆測驗	傳統團班補救教學	紙筆測驗

(二) 前測



圖 22 正式施測-前測
(線上施測 A 組)



圖 23 正式施測-前測
(線上施測 A 組)



圖 24 正式施測-前測
(紙筆測驗 B 組)



圖 25 正式施測-前測
(紙筆測驗 C 組)

(三) 補救教學

本研究意欲比較不同形式的補救教學，對學生學習成效之差異性，故根據先前規劃之分組，將 A、B 兩組共 246 人編為實驗組，採取線上適性補救教學（亦即以電腦為基礎之適性補救教學 computer-based adaptive remedial instruction, CBRI 模式），將 C 組 298 人編為對照組，採取傳統團班教學（亦即以教師為基礎之適性補救教學 teacher-based adaptive remedial instruction, TBRI 模式）。

兩大組所使用之教材完全相同，授課人員相同，授課時數相同，控制變因僅係授課工具差異（線上適性/團班教學）。並為消彌實驗組學生（線上適性）在進行補救時所衍生的後繼問題可能無法立即解決的缺點，實驗組與對照組學生（團班教學）均不採用課程中現場提問，而改以課後統一回應詢答方式。



圖 26 線上補救教學
(適性補救 A 班)



圖 27 線上補救教學
(適性補救 B 班)



圖 28 補救教學
(團班補救 C 班)



圖 29 正式施測-後測
(線上施測 A 班)

(四) 後測

不同形式的補救教學對學生學習成效之差異，表現在前測與後測之錯誤類型更正，與總體進步幅度變化上。故根據先前規劃之分組，將 A 組 125 人施以「線上適性測驗」、B 組 121 人施以「紙筆後測」、C 組 298 人施以「紙筆後測」。

對於「僅進行前測未進行後測」以及「僅進行後測未進行前測」之樣本，均視為無效樣本，予以剔除。對於「前後測均有進行，但未接受補救教學」以及「前測成績遠高於後測成績，且答題反應明顯不合邏輯」之樣本則視為不良樣本，亦暫予排除。故上述樣本數均係刪除無效樣本後之實際實驗數據。



圖 30 正式施測-後測
(紙筆測驗 B 班)



圖 31 正式施測-後測
(紙筆測驗 C 班)

肆、研究結果

一、電腦化適性測驗系統省題率效果

(一) 以適性選題方式進行施測與全測之比較

實驗組 A 組採電腦化適性測驗，當電腦化適性測驗估計結束後，測驗繼續直到學生完成整份測驗，此測驗方式為「適性全測」，將收集到的全測資料使用 5-fold cross-validation 法評估電腦化適性測驗之節省試題比率，前後測評估結果如下表 3、4 所示。

表 3 電腦化適性測驗前測閾值、預測精準度及平均施測題數表

閾值	0.026	0.031	0.036	0.041	0.046
預測精準度	0.9631	0.9663	0.9313	0.9333	0.9212
平均施測題數	16.9355	17.0336	15.9742	15.8697	15.5471
閾值	0.051	0.056	0.061	0.066	0.071
預測精準度	0.9078	0.9060	0.9025	0.8761	0.8735
平均施測題數	15.08	14.9742	14.8039	13.9716	13.8774

表 4 電腦化適性測驗後測閾值、預測精準度及平均施測題數表

閾值	0.026	0.031	0.036	0.041	0.046
預測精準度	0.980985	0.981664	0.967199	0.957895	0.943905
平均施測題數	15.49935	15.40258	14.45419	14.09806	13.51484
閾值	0.051	0.056	0.061	0.066	0.071
預測精準度	0.929508	0.917487	0.905874	0.870357	0.872122
平均施測題數	13.0671	12.64774	12.39742	11.25677	11.35871

在前測方面，當預測精準度為 96.31% 時，平均施測題數為 16.33 題，大約節

省 14.05% 的試題；當預測精準度為 90.25% 時，平均施測題數為 14.80 題，大約節省 22.11% 的試題。

後測方面，當預測精準度為 95.79% 時，平均施測題數為 14.10 題，大約節省 25.79% 的試題；當預測精準度為 90.59% 時，平均施測題數為 12.39 題，大約節省 34.79% 的試題。

此結果表示，無論前測或後測，若採用學生試題結構進行適性選題，比學生接受全測驗均可皆省試題，且後測節省效果優於前測。

二、補救教學效果比較

(一) 三組間受試者的成績變化

為了排除起點行為的差異，影響補救教學效果之分析，分別對 A、B 組及 B、C 組進行同質性分析，以前測成績校正後測成績，做組內迴歸同質性分析，並將校正後的成績做比較，分析結果如下所示：

1. A、B 組的分析結果

(1) 同質性檢定：A、B 兩組受試者間進行同質性分析，結果如表 5 所示，並未達顯著性，表示 A、B 兩組受試者間的前後測成績具同質性。

表5 A、B組受試者間效應檢定

來源	型 III 平方和	df	平均平方和	F	顯著性
分組*前測成績	79.804	1	79.804	.879	.349
誤差	21964.882	242	90.764		
校正後的總數	23520.459	245			

依變數：後測成績

R 平方 = .066 (調過後的 R 平方 = .055)

(2) A、B 組校正後之後測平均成績的差異：由於 A 組與 B 組之控制變因係表現在施測形式上 (A 組為電腦化適性測驗、B 組為紙筆測驗)，當以前測成績校正後的後測成績，得出平均 A 組比 B 組高 4.579 分結果，且差異已達顯著，表示「以電腦化適性測驗確優於紙筆測驗」。

表6 A、B組校正後的平均成績

分組	平均數	標準差	95% 信賴區間	
			下界	上界
A組	95.687 ^a	.869	93.975	97.399
B組	91.108 ^a	.884	89.367	92.849

依變數：後測成績

^a.使用下列值估計出現在模式的共變量：前測成績 = 71.50

表 7 A、B 組校正後平均成績的差異

(I)分組	(J)分組	平均差異 (I-J)	標準差	顯著性 ^a	差異的 95% 信賴區間 ^a	
					下界	上界
A 組	B 組	4.579*	1.264	.000	2.089	7.069

依變數：後測成績

根據估計的邊緣平均數而定

*.平均差異在 .05 水準是顯著的。

^a.調整多重比較：最低顯著差異（等於未調整值）。

2. B、C 組的分析結果

(1) 同質性檢定：B、C 兩組進行同質性分析，結果如表 8 所示，並未達顯著性，表示 B、C 兩組的前後測成績具同質性。

表 8 B、C 組同質性檢定

來源	型 III 平方和	df	平均平方和	F	顯著性
分組*前測成績	240.412	1	240.412	.678	.411
誤差	158553.725	447	354.706		
校正後的總數	170113.725	450			

依變數：後測成績

R 平方 = .068（調過後的 R 平方 = .062）

(2) B、C 組校正後之後測平均成績的差異：由於 B 組與 C 組之控制變因係表現在補救教學形式上（B 組為線上適性補救教材、C 組為傳統團班補救教學），當以前測成績校正後的後測成績，得出平均 B 組比 C 組高 10.161 分結果，且差異已達顯著，表示「以電腦化補救教學確優於傳統團班補救教學」。

表 9 B、C 組校正後的平均成績

分組	平均數	標準差	95% 信賴區間	
			下界	上界
B 組	90.982 ^a	1.712	87.619	94.346
C 組	80.822 ^a	1.036	78.785	82.858

表9 B、C組校正後的平均成績

分組	平均數	標準差	95% 信賴區間	
			下界	上界
B組	90.982 ^a	1.712	87.619	94.346
C組	80.822 ^a	1.036	78.785	82.858

依變數：後測成績

^a.使用下列值估計出現在模式的共變量：前測成績 = 65.45

表 10 B、C 組校正後的平均成績差異

(I) 分組	(J) 分組	平均差異 (I-J)	標準差	顯著性 ^a	差異的 95% 信賴區 ^a	
					下界	上界
B組	C組	10.161*	2.001	.000	6.229	14.093

依變數：後測成績

根據估計的邊緣平均數而定

*.平均差異在 .05 水準是顯著的。

^a.調整多重比較：最低顯著差異（等於未調整值）。

貳、三組間受試者之高低分組的成績變化

一、A、B 組高分組的分析結果

（一）同質性檢定：A、B 兩組高分組進行同質性分析，結果如表 11 所示，並未達顯著性，表示 A、B 兩組的前後測成績具同質性。

表 11 A、B 組之高分組同質性檢定

來源	型 III 平方和	df	平均平方和	F	顯著性
分組*前測成績	99.908	1	99.908	1.143	.287
誤差	10491.542	120	87.430		
校正後的總數	11286.742	123			

依變數：後測成績

R 平方 = .070（調過後的 R 平方 = .047）

（二）A、B 組之高分組校正後的後測平均成績差異

表12 A、B組之高分組校正後的平均成績

分組	平均數	標準差	95% 信賴區間
----	-----	-----	----------

			下界	上界
A組	95.619 ^a	1.443	92.763	98.475
B組	92.820 ^a	1.474	89.902	95.738

依變數：後測成績

^a.使用下列值估計出現在模式的共變量：前測成績 = 85.73

表 13 A、B 組之高分組校正後平均成績的差異

(I)分組	(J)分組	平均差異 (I-J)	標準差	顯著性 ^a	差異的 95% 信賴區 ^a 下界	上界
A 組	B 組	2.799	2.384	.243	-1.921	7.519

依變數：後測成績

根據估計的邊緣平均數而定

*. 平均差異在 .05 水準是顯著的。

^a. 調整多重比較：最低顯著差異（等於未調整值）。

二、A、B 組低分組的分析結果

（一）同質性檢定：A、B 兩組低分組進行同質性分析，結果如表 14 所示，並未達顯著性，表示 A、B 組之低分組的前後測成績具同質性。

表 14 A、B 組之低分組同質性檢定

來源	型 III 平方和	df	平均平方和	F	顯著性
分組*前測成績	.590	1	.590	.006	.937
誤差	11206.137	118	94.967		
校正後的總數	12070.893	121			

依變數：後測成績

R 平方 = .072（調過後的 R 平方 = .048）

（二）A、B 組之低分組校正後的後測平均成績差異

表 15 A、B 組之低分組校正後的平均成績

分組	平均數	標準差	95% 信賴區間	
			下界	上界
A組	95.322 ^a	1.254	92.838	97.806
B組	89.817 ^a	1.276	87.291	92.343

依變數：後測成績

^a.使用下列值估計出現在模式的共變量：前測成績 = 64.56

表 16 A、B 組之低分組校正後平均成績的差異

(I)分組	(J)分組	平均差異 (I-J)	標準差	顯著性 ^a	差異的 95% 信賴區 ^a	
					下界	上界
A 組	B 組	5.505*	1.820	.003	1.901	9.109

依變數：後測成績

根據估計的邊緣平均數而定

*.平均差異在 .05 水準是顯著的。

^a.調整多重比較：最低顯著差異（等於未調整值）。

三、B、C 組高分組的分析結果

（一）同質性檢定：B、C 組兩組高分組進行同質性分析，結果如表 17 所示，並未達顯著性，表示 B、C 組之高分組的前後測成績具同質性。

表 17 B、C 組之高分組的同質性檢定

來源	型 III 平方和	df	平均平方和	F	顯著性
分組*前測成績	811.307	1	811.307	2.764	.098
誤差	65458.191	223	293.534		
校正後的總數	74365.956	226			

依變數：後測成績

R 平方 = .120（調過後的 R 平方 = .108）

（二）B、C 組之高分組校正後的後測平均成績差異

表 18 B、C 組之高分組校正後的平均成績

分組	平均數	標準差	95% 信賴區間	
			下界	上界
B 組	91.899 ^a	2.202	87.559	96.239
C 組	83.447 ^a	1.335	80.816	86.078

依變數：後測成績

^a.使用下列值估計出現在模式的共變量：前測成績 = 78.99

表 19 B、C 組之高分組校正後平均成績的差異

(I)分組	(J)分組	平均差異 (I-J)	標準差	顯著性 ^a	差異的 95% 信賴區 ^a	
					下界	上界
B 組	C 組	8.452*	2.575	.001	3.376	13.527

依變數：後測成績

根據估計的邊緣平均數而定

*.平均差異在 .05 水準是顯著的。

(I)分組	(J)分組	平均差異 (I-J)	標準差	顯著性 ^a	差異的 95% 信賴區 ^a	
					下界	上界
B組	C組	8.452*	2.575	.001	3.376	13.527

依變數：後測成績

根據估計的邊緣平均數而定

*.平均差異在 .05 水準是顯著的。

^a.調整多重比較：最低顯著差異（等於未調整值）。

四、B、C 組低分組的分析結果

（一）同質性檢定：B、C 組兩組低分組進行同質性分析，結果如表 20 所示，並未達顯著性，表示 B、C 組之低分組的前後測成績具同質性。

表 20 B、C 組之低分組的同質性檢定

來源	型 III 平方和	df	平均平方和	F	顯著性
分組*前測成績	166.616	1	166.616	.423	.516
誤差	86572.955	220	393.513		
校正後的總數	93594.839	223			

依變數：後測成績

R 平方 = .075（調過後的 R 平方 = .062）

（二）B、C 組之高低分組校正後的後測平均成績差異

表 21 B、C 組之低分組校正後的平均成績

分組	平均數	標準差	95% 信賴區間	
			下界	上界
B組	89.939 ^a	2.558	84.898	94.979
C組	78.205 ^a	1.547	75.157	81.254

依變數：後測成績

^a.使用下列值估計出現在模式的共變量：前測成績 = 51.74

表 22 B、C 組低分組校正後的平均成績差異

(I)分組	(J)分組	平均差異 (I-J)	標準差	顯著性 ^a	差異的 95% 信賴區 ^a	
					下界	上界
B組	C組	11.733*	2.989	.000	5.842	17.624

依變數：後測成績

根據估計的邊緣平均數而定

*.平均差異在 .05 水準是顯著的。

^a.調整多重比較：最低顯著差異（等於未調整值）。

五、各組間之高低分組校準後的後測成績差異

整體評估各組組間高低分組後測校準後成績差異量，統整為表23，發現A組與B組之低分群，B組與C組之高分群及低分群，均顯現明顯之成績差異，且均係前組高於後組，表示「以電腦化適性測驗達成之後測成績，確實優於以紙筆測驗達成之後測成績」，以及「以線上適性補救教學達成之後測成績，確實優於以傳統團班補救教學達成之後測成績」。

表 23 各組組間高低分組後測校準後成績差異量

組別 (I、J)	高低分組	成績差異量 (I-J)	是否達顯著
A、B 組	高分組	2.799	否
	低分組	5.505	是
B、C 組	高分組	8.452	是
	低分組	11.733	是

儘管A組與B組之高分群並未達顯著差異，唯已達2.799之差值，且係前組高於後組，可謂雖不中亦不遠矣。此結果亦顯示以電腦化適性施測的方式對於高分組的學生較無差異，但對於低分組的受試者來說，可能因為施測工具引起學習興趣，使得以電腦化適性施測的方式的學生後測成績進步量顯著優於以紙筆方式施測的組別。

(二) 各組內前後測的成績差異

1. 各組前後測成績差異

表24 各組前後測的成績差異

組別	成對變數差異 (前測-後測)					顯著性 (雙尾)
	平均數	標準差	平均數的 標準差	差異的 95% 信賴區間 下界	上界	
A 組	-18.456	24.262	2.170	-22.751	-14.161	.000
B 組	-20.169	24.756	2.123	-24.367	-15.971	.000
C 組	-20.641	20.343	1.178	-22.960	-18.322	.000

整體評估各組組內前後測成績進步量，發現均顯現明顯之成績差異，且均係後測高於前測，表示各組各群經補救教學後，成績大抵有所提升。且以線上適性補救教學達成之後測進步成績，確實優於以傳統團班補救教學達成之後測進步成績。

伍、各組內之高低分組前後測成績的差異

一、A 組之高分組前後測成績的差異

表25 A組之高分組前後測成績的差異

項目	成對變數差異					顯著性 (雙尾)
	平均數	標準差	平均數的 標準差	差異的	95% 信賴區間	
				下界	上界	
前測-後測	-4.333	5.708	.719	-5.771	-2.896	.000

二、A 組之低分組前後測成績的差異

表26 A組之低分組前後測成績的差異

項目	成對變數差異					顯著性 (雙尾)
	平均數	標準差	平均數的 標準差	差異的	95% 信賴區間	
				下界	上界	
前測-後測	-32.806	27.353	3.474	-39.753	-25.860	.000

三、B 組之高分組前後測成績的差異

表27 B組之高分組前後測成績的差異

項目	成對變數差異					顯著性 (雙尾)
	平均數	標準差	平均數的 標準差	差異的	95% 信賴區間	
				下界	上界	
前測-後測	-12.836	14.424	1.847	-16.530	-9.142	.000

四、B 組之低分組前後測成績的差異

表28 B組之低分組前後測成績的差異

項目	成對變數差異					顯著性 (雙尾)
	平均數	標準差	平均數的 標準差	差異的	95% 信賴區間	
				下界	上界	
前測-後測	-38.417	17.151	2.214	-42.847	-33.986	.000

五、C 組之高分組前後測成績的差異

表29 C組之高分組前後測成績的差異

項目	成對變數差異					顯著性 (雙尾)
	平均數	標準差	平均數的 標準差	差異的	95% 信賴區間	
				下界	上界	
前測-後測	-4.488	18.686	1.450	-7.352	-1.624	.002

六、C 組之低分組前後測成績的差異

表30 C組之低分組前後測成績的差異

項目	成對變數差異					顯著性 (雙尾)
	平均數	標準差	平均數的 標準差	95% 信賴區間		
				下界	上界	
前測-後測	-26.384	26.713	2.086	-30.503	-22.265	.000

七、各組之高低分組前後測成績的進步量

整體評估各組組內高低分組後測成績差異量，統整為表31，發現各組之高低分群，均顯現明顯之成績差異，且所有分組分群均係後測高於前測，表示各組各群經補救教學後，成績大抵有所提升，意味此實驗所進行之補救教學成效良好。同時發現對各組低分群之補救效果至為可觀（高達20~30分），且以線上適性補救教學達成之後測進步成績，確實優於以傳統團班補救教學達成之後測進步成績。

表 31 各組前後測高低分組成績進步量

組別	高低分組	成績進步量（前測-後測）	是否達顯著
A 組	高分組	-4.333	是
	低分組	-32.806	是
B 組	高分組	-12.836	是
	低分組	-38.417	是
C 組	高分組	-4.488	是
	低分組	-26.384	是

陸、電腦化測驗重點通過率的變化

整體評估A組前後測技能通過率比較，如表32，發現每一個技能類型均獲得進步（19個/19個重點），甚至有達到100%的通過率（重點1、4、6、8、9、11），其他重點也有亮眼成果（重點7由24%進至98%、重點12由50%進至94%），而絕大多數（17個/19個重點）達90%~100%。

至於B組前後測技能通過率比較，如表33，則發現大體上呈現進步（15個/19個重點），且進步幅度甚高（重點3由69%進至99%、重點7由26%進至99%、重點10由56%進至93%、重點12由66%進至96%、重點13由50%進至82%、重點14由55%進至94%、重點15由34%進至90%、重點16由52%進至90%、重點17由46%進至95%、重點18由53%進至96%），但仍有部分（4個/19個重點）主題出現下滑（重點2、5、6、11）。

再看C組前後測技能通過率比較，如表34，則發現大致呈現進步（14個/19個重點），也有引人注目的進步幅度（重點2由63%進至96%、重點7由24%進至95%、重點9由67%進至90%、重點12由58%進至90%、重點13由52%進至93%、重點14由57%進至94%、重點15由32%進至87%、重點16由47%進至92%、重點17由51%進至89%），但亦有部分（5個/19個重點）主題出現下滑（重點1、4、5、6、8）。

關於部分主題通過率下滑之原因，基於以電腦適性施測之A組，其前後測通過率進步幅度明顯高於以紙筆施測之B、C組，有可能係以電腦施測較具新鮮感，

學生答題較用心謹慎，至於紙筆施測組則因複本施測使學生答題產生疲憊感，故不良樣本數明顯增加，前後測通過率進步幅度無法像電腦施測組般顯著。

綜合上述結果，顯現以「電腦化適性測驗+電腦化補救教學」確產生令人驚嘆的教學效果，遠優於「紙筆測驗+傳統團班補救教學」。

表 32 A 組前後測技能通過率比較表

技能類型	診斷結果：通過率	
	卷 1	卷 2
【重點 1】矩陣的定義	96%	100%
【重點 2】轉置矩陣	87%	99%
【重點 3】對稱反對稱矩陣	69%	96%
【重點 4】矩陣的加減	96%	100%
【重點 5】矩陣的係數積	93%	97%
【重點 6】矩陣的乘法	95%	100%
【重點 7】馬可夫矩陣	24%	98%
【重點 8】高斯消去法	94%	100%
【重點 9】矩陣的三角化	92%	100%
【重點 10】矩陣的列運算與單位矩陣變換	91%	99%
【重點 11】反矩陣	89%	100%
【重點 12】冪矩陣	50%	94%
【重點 13】旋轉矩陣	81%	99%
【重點 14】鏡射矩陣	81%	99%

表 32 A 組前後測技能通過率比較表（續）

技能類型	診斷結果：通過率	
	卷 1	卷 2
【重點 15】伸縮矩陣	79%	98%
【重點 16】推移矩陣	81%	99%
【重點 17】矩陣的綜合線性變換	78%	97%
【重點 18】矩陣的拆解	53%	75%
【重點 19】線性變換與面積變換	46%	71%

表 33 B 組前後測技能通過率比較表

技能類型	診斷結果：通過率	
	卷 1	卷 2
【重點 1】矩陣的定義	81%	93%
【重點 2】轉置矩陣	96%	91%
【重點 3】對稱反對稱矩陣	69%	99%
【重點 4】矩陣的加減	72%	94%

【重點 5】矩陣的係數積	97%	95%
【重點 6】矩陣的乘法	95%	66%
【重點 7】馬可夫矩陣	26%	99%
【重點 8】高斯消去法	87%	89%
【重點 9】矩陣的三角化	82%	93%
【重點 10】矩陣的列運算與單位矩陣變換	56%	93%
【重點 11】反矩陣	93%	87%
【重點 12】冪矩陣	66%	96%
【重點 13】旋轉矩陣	50%	82%
【重點 14】鏡射矩陣	55%	94%
【重點 15】伸縮矩陣	34%	90%
【重點 16】推移矩陣	52%	90%
【重點 17】矩陣的綜合線性變換	46%	95%
【重點 18】矩陣的拆解	53%	96%
【重點 19】線性變換與面積變換	57%	81%

表 34 C 組前後測技能通過率比較表

技能類型	診斷結果：通過率	
	卷 1	卷 2
【重點 1】矩陣的定義	92%	84%
【重點 2】轉置矩陣	63%	96%

表 34 C 組前後測技能通過率比較表（續）

技能類型	診斷結果：通過率	
	卷 1	卷 2
【重點 3】對稱反對稱矩陣	72%	97%
【重點 4】矩陣的加減	95%	85%
【重點 5】矩陣的係數積	97%	93%
【重點 6】矩陣的乘法	95%	51%
【重點 7】馬可夫矩陣	24%	95%
【重點 8】高斯消去法	84%	81%
【重點 9】矩陣的三角化	67%	90%
【重點 10】矩陣的列運算與單位矩陣變換	83%	87%
【重點 11】反矩陣	80%	88%
【重點 12】冪矩陣	58%	90%
【重點 13】旋轉矩陣	52%	93%
【重點 14】鏡射矩陣	57%	94%
【重點 15】伸縮矩陣	32%	87%

【重點 16】推移矩陣	47%	92%
【重點 17】矩陣的綜合線性變換	51%	89%
【重點 18】矩陣的拆解	56%	69%
【重點 19】線性變換與面積變換	52%	74%

伍、結論

根據研究目的與實驗結果，研究結論如下所述：

根據課綱，歷經專家會議之審檢，建立了高中「矩陣」單元的專家知識結構，並遵守Prosser（1974）的建議，建構了每個概念至少包含10個試題的對應題庫，已能負荷電腦化適性測驗之需求。進而建置電腦化適性選填題測驗診斷系統，此系統之介面簡單且易於操作，並適用電腦或行動載具進行施測，藉以減少傳統畫卡式作答與閱卷所造成之人時壓力。

關乎本電腦化適性選填題測驗診斷系統之省題率效能，經評估結果顯示：在前測方面，當預測精準度為96.31%時，平均施測題數為16.33題，大約節省14.05%的試題；當預測精準度為90.25%時，平均施測題數為14.80題，大約節省22.11%的試題。後測方面，當預測精準度為95.79%時，平均施測題數為14.10題，大約節省25.79%的試題；當預測精準度為90.59%時，平均施測題數為12.39題，大約節省34.79%的試題。表示無論前測或後測，足堪確認「採用學生試題結構進行適性選題，比學生接受全測驗可節省試題」。

同時此適性測驗系統，可以診斷高中生矩陣單元學習狀況，鑑別學生個別盲點，輔以「針對性補救教學結構」，進行適性補救教學，並獲得：以前測成績校正後的後測成績，電腦化適性測驗組比紙筆測驗組平均高4.579分結果，且在顯著水準為0.00下，統計上達顯著差異。各組前後測成績差異後測的平均成績高於前測的平均成績逾20分，且在顯著水準為0.00下，統計上達顯著差異。以及以前測成績校正後的後測成績，線上適性補救比傳統團班補救成績達10.161分，且在顯著水準為0.00下，統計上達顯著差異。足證「經補救教學後各組前後測成績進步達顯著差異」、「電腦化補救教學優於傳統團班補救教學」，提升高中教學成效可以期待。

陸、未來研究方向

由於本研究係在KSAT系統下，率先以高中生為實驗主體之實徵資料探討，對於實驗施行過程蓋無前例可循，故尚有諸多可改進空間。茲分述如下：

一、以電腦適性施測之A組，其前後測通過率進步幅度明顯高於以紙筆施測之B、C組，有可能係以電腦施測較具新鮮感，學生答題較用心謹慎，至於紙筆施測組則因複本施測使學生答題產生疲憊感，故不良樣本數明顯增加，前後測通過率進步幅度無法像電腦施測組般顯著。未來在實驗設計上，尤其是複本之新穎度上，應有更佳規劃，俾降低不良樣本比例，提高實驗準確度。

二、本研究可供進一步探究目標預估為：進一步完成高中三個年級完整的專家知識結構與相對應題庫。並建置可由教學者自由調整調用的專家知識結構，與對應電腦化適性題庫系統。進而確認KSAT系統在高中各單元的強韌性。

參考文獻

中文部分

- 王惠美 (2002)。適性化網頁教學系統之學習迷失原因診斷與回饋—以英語介係詞為例。中華大學資訊工程學系碩士論文，未出版，新竹市。
- 江淑卿、郭生玉 (1997)。不同學習過程的概念構圖策略對促進知識結構專家化與理解能力之效果研究。師大學報，42，1-16。
- 汪端正 (2008)。適性診斷測驗與數位個別指導教材之研發—以國小六年級質數與合數單元為例。國立臺中教育大學教育測驗統計研究所碩士論文，未出版，臺中市。
- 余民寧 (2009)。試題反應理論 (IRT) 及其應用。心理出版社，臺北市。
- 吳玫君、施淑娟、許天維、陳淑勤 (2008)。國小五年級數學領域概數與估算單元數位元個別指導模式之研發。第四屆臺灣數位學習發展研討會，TWELF 2008。2008年5月17日，臺中：國立臺中教育大學。
- 吳慧珉 (2006)。以知識結構為基礎之適性測驗選題策略強韌性探究。測驗統計年刊，14 (2)，1-16。
- 林婉星 (2008)。以知識結構為基礎之電腦化適性數學測驗線上更新結構機制之研發。國立臺中教育大學教育測驗統計研究所碩士論文，未出版，臺中市。
- 林立敏、白曉珊、郭伯臣、劉育隆 (2007)。數位個別指導教材研發與適性補救教學模式之研究—以國小五年級數學「因數與倍數」單元為例。網際網路技術學刊，8 (2)，191-197。
- 俞克斌、施淑娟、郭伯臣、鄭俊彥 (2012)。建置基於DINA模式之高中數學電腦化診斷測驗系統初探。2012全球華人計算機教育應用大會工作坊，GCCCE2012。2012年5月28日-6月1日，墾丁：福華飯店。
- 洪素敏 (2004)。國小五年級學童分數迷思概念補救教學之研究。國立嘉義大學數學教育研究所碩士論文，未出版，嘉義市。
- 涂金堂、林佳蓉 (2000)。如何協助學生解決數學應用問題。復文出版社，高雄市。
- 徐發興 (2011)。學測成績通知單樣式暨選擇 (填) 題計分方式。選才電子報，198。取自：<http://www.ceec.edu.tw/CeecMag/Articles/198/198-04.htm>
- 陳世銘、李淑娟、施淑娟、吳任婕 (2008)。符合SCORM 2004與試題結構理論之電腦化適性測驗及補救教學系統研發—以國小數學領域五年級能力指標幾何為例。測驗統計年刊，16 (2)，31-58。
- 教育部 (2010)。普通高級中學課程綱要 (97年版，99 學年度高一新生適用)。教育部中等教育司。取自：
http://www.edu.tw/high-school/content.aspx?site_content_sn=8403
- 張雅玲 (2010)。結構化教學與補救教學成效探討—以「小數的加減」為例。國立臺中教育大學教育測驗統計研究所碩士論文，未出版，臺中市。
- 張新仁 (1993)。不同寫作能力的國小學童寫作歷程之研究。國科會專題研究計

- 畫成果報告（編號：NSC82-0301-H017-001），未出版。臺北市：中華民國行政院國家科學委員會。
- 黃碧雲（2005）。以能力指標結構為基礎的電腦適性測驗編製及動畫補救教學之應用—以國小數學領域四年級能力指標。國立臺中師範學院數學教育學系碩士論文，未出版，臺中市。
- 黃珮璇、王暄博、郭伯臣、劉湘川（2006）。國小數學科電腦化適性診斷測驗強韌性探究。2006年電腦與網路科技在教育上的應用研討會。2006年3月23-24日，台中：國立新竹教育大學。
- 郭伯臣（2003）。國小數學科電腦化適性診斷測驗（I）。行政院國家科學委員會專題研究成果報告（編號：NSC91-2520-S-142-001-），未出版，臺北市：中華民國行政院國家科學委員會。
- 郭伯臣（2004）。國小數學科電腦化適性診斷測驗（II）。行政院國家科學委員會專題研究成果報告（編號：NSC92-2521-S-142-003-），未出版，臺北市：中華民國行政院國家科學委員會。
- 郭伯臣（2005）。國小數學科電腦化適性診斷測驗（III）。行政院國家科學委員會專題研究成果報告（編號：NSC93-2521-S-142-004-），未出版，臺北市：中華民國行政院國家科學委員會。
- 郭伯臣、謝友振、張峻豪、蔡坤穎（2005）。以結構理論為基礎之適性測驗與適性補救教學線上系統。臺灣數位學習發展研討會，臺北：國立臺灣師範大學。
- 曾彥鈞（2007）。以知識結構為基礎的適性診斷測驗系統及降低猜測機制之研發。國立臺中教育大學數學教育學系碩士論文，未出版，臺中市。
- 楊智為、劉育隆、楊晉民、曾彥鈞（2006）。結合試題順序理論與貝氏網路之電腦適性測驗演算法之探究。第十一屆人工智慧與應用研討會。2006年12月15日-16日，高雄：高雄應用科技大學。
- 趙琬津（2005）。數位個別指導模式與教材之研發-以「三角型」單元為例。國立台中教育大學數學教育研究所碩士論文，未出版，臺中市。
- 劉安哲（2007）。使用自動化概念結構建製法於貝氏網路學生模式建立及成效分析—以國小數學領域小數加減為例。亞洲大學資訊工程學系碩士論文，未出版，臺中市。
- 劉易昌（2008）。以知識結構為基礎之適性測驗曝光率管控法研發。亞洲大學資訊工程學系碩士論文，未出版，臺中市。
- 劉育隆、曾筱倩、郭伯臣（2006）。以知識結構為基礎之適性測驗系統建置。測驗統計年刊，14（2），17-35。

外文部分

- Appleby, J., Samules, P. & Treasure-Jones, T. (1997). Diagnosis a knowledge-based diagnostic test of basic mathematical skills. *Computers & Education*, 28(2), 113-131.

- Bower, G. H. (1972). Mental imagery and associative learning. In L. Gregg (Ed.), *Cognition in learning and memory*, 51-88.
- Brown, J. S. & Burton, R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, 2, 155-192.
- Chang, K. E., Liu, S. H. & Chen, S. W. (1998). A testing system for diagnosing misconceptions in DC electric circuits. *Computers & Education*, 31, 195-210.
- Chi, M. T. H., Glaser, R. & Rees, E. (1982). Expertise in problem solving. In R.J. Sternberg (Ed.), *Advances in psychology of human intelligence*, 1, 7-75. Hillsdale, NJ: Erlbaum.
- Hall, R. (1988). *Knowledge maps and the presentation of related information domains*. Unpublished doctoral dissertation, Texas Christian University, Fort Worth, TX.
- Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, 208, 1335-1342.
- Prosser, F. (1974). Item banking. In G. Lippey (Ed.), *Computer-assisted test construction*, 29-66. Englewood Cliffs, NJ: Educational Technology.
- Reckase, M. D. (1981). Tailored testing, measurement problems and latent trait theory. Paper presented at *the annual meeting of the National Council for Measurement in Education*, Los Angeles.
- Rewey, K. L., Dansereau, D. F., Skaggs, L. P., Hall, R. H. & Pitre, U. (1989). Effects of scripted cooperation and knowledge maps on the processing of technical material. *Journal of Educational Psychology*, 81(4), 604-609.
- Shavelson, R. J. (1972). Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Journal of Educational Psychology*, 63, 225-234.
- Shavelson, R. J., & Stanton, G. C. (1975). Construct validation: Methodology and application to three measures of cognitive structure. *Journal of Educational Measurement*, 12(2), 67-85.
- Skaggs, L. P. (1988). *The effects of knowledge maps and pictures on the acquisition of scientific information*. Unpublished doctoral dissertation, Texas Christian University, Fort Worth, TX.
- Thro, M. P. (1978). Relationships between associative and content structure of physics concepts. *Journal of Educational Psychology*, 70(6), 971-978.
- VanLehn, K.(1988). *Student Modeling*. In M.C. Polson and J.J Richardson. Foundations of Intelligent Tutoring Systems. Hillsdale, NJ: Lawrence Erlbaum Publishers.
- Wainer, H. (1990). *Computerized adaptive testing: A primer*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wainer, H. (2000). *Computerized adaptive testing: A primer* (2nd ed). Hillsdale, NJ:

Lawrence Erlbaum Publishers.

Wenger, E. (1987). *Artificial Intelligence and Tutoring Systems: Computational and cognitive approaches to the communication of knowledge*. Los Altos, California: Morgan Kaufmann Publishers.