

網路適性測驗在特殊教育之應用

陳錫鴻

教育部中部辦公室

e-1303@mail.tpde.edu.tw

郭伯臣

國立台中教育大學

kbc@mail.ntctc.edu.tw

劉湘川

亞洲大學

lhc@thmu.edu.tw

摘要

目前國內外以試題反應理論(Item Response Theory)探討智能障礙學生的表現情形與潛在特質(latent traits)或能力(abilities)間關係的研究並不多，本論文以由師大心測中心發展國中基本學力測驗的關鍵核心技術IRT為基礎，將試題反應理論及電腦化適性測驗評量原理應用於特殊教育領域，希望能發展出一套適合身心障礙學生能力評估且兼具科學理論基礎之可行模式。

本研究探討以高職特教班課程綱要編製實用數學能力重要指標及試題，運用試題反應理論三參數對數模式分析，以了解高職特教班學生實用數學程度。最後建置高職特教班實用數學鑑定適性測驗系統，以提供個別學生最大訊息量試題為適性選題策略，由網際網路遠距線上施測，根據受試者各答題組型，決定其能力可能落點，並藉由電腦強大運算能力，即時選題施測。

研究結果發現：電腦化適性測驗符合特殊教育「個別化」、「適性化」精神：本系統每個學生線上施測的題次均不同，以最符合個別學生能力試題進行評量。施測結果節省試題及作答時間平均達23.56%，且仍能準確評估個別智能障礙類學生數學能力。

關鍵詞：試題反應理論、電腦化適性測驗、高職特教班、特殊教育。

Abstract

Referring to the Item Response Theory (IRT), there are actually few researches of the relation between the performance and the latent traits or the abilities of mentally retarded students. The prime goal of this study is to introduce IRT and Computerized Adaptive Test (CAT) to special education, hoping to develop a feasible way to evaluate mentally retarded students' ability with a scientific method.

With the better understanding of students' ability in the subject, we can then use IRT as a basis to construct a useful CAT system on the network. By the powerful calculation of the computer, the CAT system can immediately assigns question to target students by their item response pattern. The pool of questions consider offering the largest amount of information to subject.

The CAT meets the 'Individualized' and 'Adaptive' needs of what is required in special education. By the system, each student who takes the

online tests will face different series of questions. In line with student's individual ability, the amount of questions on the test will be given accordingly. The result of examining can save the examination questions and answering time up to 23.56% in average. Moreover, we also can exactly evaluate specific mentally retarded student's mathematics ability accurately.

The results of research are: The CAT meets the 'Individualized' and 'Adaptive' needs of what is required in special education. By the system, each student who takes the online tests will face different series of questions. In line with student's individual ability, the amount of questions on the test will be given accordingly. The result of examining can save the examination questions and answering time up to 23.56% in average. Moreover, we also can exactly evaluate specific mentally retarded student's mathematics ability accurately.

Keywords: item response theory, computerized adaptive test, special education.

1. 前言

扶助弱勢群族教育為政府宣示施政重點主軸之一，為身心障礙學生的抽象思考能力較為薄弱，網路資源運用教學正可以提供具體化、生活化、趣味化以及實用化的學習經驗(朱經明，民86)。所以洪榮昭(民88)指出對於特殊教育而言，應用電腦在低學習能力程度者比高學習能力學習者效果好。所以就特殊教育各學科教學領域如實用語文、實用數學、生活教育、社會適應...等均可透過電腦網路獲得許多教學的素材與教學策略。結合電腦網路資源的教學已是目前的教學新趨勢，電腦網路科技運用於特殊教育將由傳統的紙上作業，漸次提升到電腦網路多媒體的歷程，同時以特教學生為主的網路學習環境及網路多媒體資源教學的新一代學習策略，這將會是一項很大的突破(王崇懋，民93)。

目前國內外以試題反應理論在特殊教育應用研究並不多見，本文嘗試將試題反應理論及電腦化適性測驗評量原理應用於特殊教育領域，而以具備特殊教育「個別化」、「適性化」精神之電腦測驗方式，對於輕度智能障礙過動及缺乏答題耐心的特性(朱經明，民86)，探討減少施測題數或時間能有效作答，提高評量的效度，採數學能力鑑定電腦化適性測驗之研究，解決有效鑑定智能障礙學生實際起點行為及能力的困難。

2. 文獻探討

2.1 試題反應理論

當代測驗理論以試題反應理論 (Item Response Theory, 簡稱 IRT) 為其中心架構, IRT 就是以數學式表示受試者能力與試題難易度、鑑別度及猜測度等參數間的關係, 而這個數學關係即是機率, 以受試者的答對機率表示縱軸, 受試者的表現情形與這組潛在能力特質間的關係, 可透過一條連續性遞增的函數來加以詮釋, 這個函數便叫作試題特徵曲線 (item characteristic curve, 簡寫為 ICC)

代表 IRT 中的常用的試題反應模式可分為單參數對數模式、雙參數對數模式和三參數對數模式等, 它們按照難易度、鑑別度、猜測度等試題參數的多寡涵義分別代表著: 某一試題的正確反應機率除了受考生或受試者的能力值決定外, 並且分別受一個參數 (即難度)、二個參數 (即難度和鑑別度)、或三個參數 (即難度、鑑別度、和猜測度) 的試題參數所決定, 其正確反應機率值亦介於 0 與 1 之間。

IRT 參數模式建立在下列基本假設之上 (Hambleton, 1985; Swaminathan, 1983; Reckase, 1979; 余民寧, 民 86), 包含 1. 測驗特質或能力單向性。2. 試題間局部獨立性。3. 知道--正確反映作答假設。4. 非速度性作答。5. 模式符合性檢定。

2.2 試題反應理論優點

IRT 有別於古典測驗理論, 係建立在數理統計基礎, 其立論與假設嚴謹, 因此具有下列幾項優點:

1. 試題參數 (item parameters) (如: 難度、鑑別度、猜測度等), 是一種不受樣本影響 (sample-free) 的指標。
2. 提供個別差異的測量誤差指標, 因此能夠精確推估受試者的能力估計值。
3. 提出以試題訊息量 (item information) 及測驗訊息量 (test information) 的概念。
4. 同時考慮受試者的反應組型與試題參數等特性, 因此在估計個人能力時, 除了能夠提供一個較精確的估計值外, 對於原始得分相同的受試者, 也往往給予不同的能力估計值。
5. 採用的模式適合度考驗值, 可以提供考驗模式與資料間之適合度、受試者的反應。

2.2.1 試題訊息函數

試題訊息函數 (item information function, IIF) 為描述試題或測驗、挑選測驗試題、以及比較測驗的相對效能, 作為建立、分析、與診斷測驗的主要參考依據。函數的定義如公式 (2.1):

$$I(\theta) = \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)} \quad i=1,2,\dots,n \quad (2.1)$$

其中 $I_i(\theta)$ 代表試題 i 在能力為 θ 上所提供的訊息

$P_i(\theta)$ 為能力 θ 在試題 i 上的試題反應函數

$Q_i(\theta) = 1 - P_i(\theta)$; $P'_i(\theta)$ 為在 θ 點上的 $P_i(\theta)$ 值的導數。

試題訊息函數可以應用到前面所談到的單

參、雙參、與三參數對數試題反應模式, 這些模式都適合用於二分法計分 (dichotomously scored) 的測驗資料。Birnbaum (1968) 指出, 某個試題所提供的最大訊息量, 剛好出現在能力估計參數 θ 最可能出現的點上。

2.2.2 測驗訊息函數

根據 Birnbaum (1968) 的推導, 一份測驗的訊息函數 $I(\theta)$ 是指它在某一個 θ 值上所提供各試題訊息函數之總和, 記如公式 (2.2):

$$I(\theta) = \sum_{i=1}^n I_i(\theta) \quad (2.2)$$

測驗訊息函數最大的特性是可加性, 隨著試題數越多, 訊息量也隨之增大, 而試題訊息量之間則無關, 彼此為獨立。而在 θ 值上的測驗訊息量與該能力的估計標準誤成平方根反比, 其如下公式 (2.3) 表示:

$$SE(\hat{\theta}) = \frac{1}{\sqrt{I(\theta)}} \quad (2.3)$$

其中 $SE(\hat{\theta})$, 稱作估計標準誤 (standard error of estimation)。當 $I(\theta)$ 值達到最大時, $SE(\hat{\theta})$ 值便達到最小, 也就是說該 θ 值的最大近似估計值的估計誤差達到最小, 亦即此時的 θ 的最大近似估計值最精確。

2.3 電腦適性測驗

電腦化適性測驗係電腦測驗結合試題反應理論及適性測驗之優點發展的測驗, 測驗系統可依受試者能力水準的估計值自動選擇符合受試者能力之試題加以施測, 並依據受試者在新呈現的題目答題的對與錯情形, 重新計算受試者的能力估計值, 並據以選取下一題目, 直到受試者能力值之估計已相當地精確為止。這種測驗方式, 每位受試者被施測之題目不同, 施測時間及題數也可以不同。可以在較少的題數達到事先設定的某一精確水準, 精確地估算出受試者的能力。

將一般 CAT 實施關鍵技術及程序分起始題目、選題策略與終止條件作一簡介與探討:

(1). 起始題目設定

電腦適性測驗在測驗之始, 受試者之能力高低未能得知, 但必須選擇測驗起始題目提供施測作答, 藉以收集足夠判斷受試者能力的答題組別資料。常用於起始題的決定方式, 有以下幾種 (王寶墉, 民 84; 陳麗如, 民 87; 陳俊宏, 民 93): 1. 中等難度題目。2. 由受試者現有之特性資料或其他參考測驗結果決定。3. 自由選題。4. 隨機選題。

(2). 選題策略

選題策略乃電腦適性測驗中重要的核心要素之一, 根據不同的選題策略會導致不同的測驗效率, 常用的選題方法如下: 1. 最接近難度法 (Reckase, 1973; Urry, 1970; Weiss,

1974)。2. 最接近偏移難度法(Birnbaum, 1968)。3. 最大訊息選題法(Lord, 1980)。

(3). 終止條件

電腦化適性測驗施測題目與題數的設計因受試目的及對象均不同，而實務上測驗終止可運用下列方式(陳麗如, 民 87; 陳新豐, 民 88): 1. 設定最大施測題數。2. 當受試者的能力估計標準差低於預設標準時，測驗即終止。3. 能力的估計已穩定，再做題目已經沒有幫助時，即終止測驗。

2.4 特殊教育電腦化需求

我國 86 年新修訂特殊教育法將個別化教育計畫(individualized education program, IEP)納入法條，規定 IEP 應評量學生下列能力，包括認知能力、溝通能力、行動能力、情緒、人際關係、感官功能、健康狀況、生活自理能力、國文、數學等學業能力之現況。這些能力的評估可以透過生理檢查、標準化測驗或非標準化測驗等方式完成(林寶貴, 民 89)。惟對 IEP 的擬定過程繁瑣，近年來國內也有以下改進的意見：一、發展簡化 IEP 表格(李翠玲, 民 88)、二、建置 IEP 電腦化之學習目標資料庫需求(王華沛, 民 91; 朱經明, 民 86)，以減少教師負擔。

發展 IEP 電腦化結果的確可以節省老師們撰寫學習目標的時間，確也導致學習目標「範本」過度浮濫使用，忽略了特教所強調的「個別化」。追根究底，特教教學的實施並未能根據學生實際起點行為及能力，其困難點在於國內迄今亦尚缺乏可以提供輔助評估障礙學生潛在能力之科學工具。

電腦適性測驗是利用電腦網路來實施測驗，處理選題、計分、能力估計等問題的一種測驗方式，電腦適性測驗比傳統測驗需要施測較少的題目，就能有效估計能力值。因此，對於輕度智能障礙過動及缺乏答題耐心的特性(朱經明, 民 86)，減少施測題數或時間反而更能有效作答，精準提高評量的效度，數學能力鑑定電腦化適性測驗之研究，恰能解決有效鑑定智能障礙學生實際起點行為及能力的困難。因此，開發鑑定電腦化適性測驗處理選題模式，運用題目訊息量選題機制，減少施測題目並能有效估算學生學習能力，運用於特殊教育領域之教學評量有其必要性。

3. 研究方法

3.1 研究方法與步驟

本研究係依據發展試題及命題流程，編製符合試題反應理論(IRT)模式之試卷，並得到項目參數估計值及試題訊息函數，建置一份於網際網路使用之電腦化適性測驗，利用電腦強大運算功能，根據線上受試者答題組型，並證明運用選題機制確能在固定答題或減少施測題數下，達到預測受試者能力特質之效果。其研究架構如圖 3.1.1：

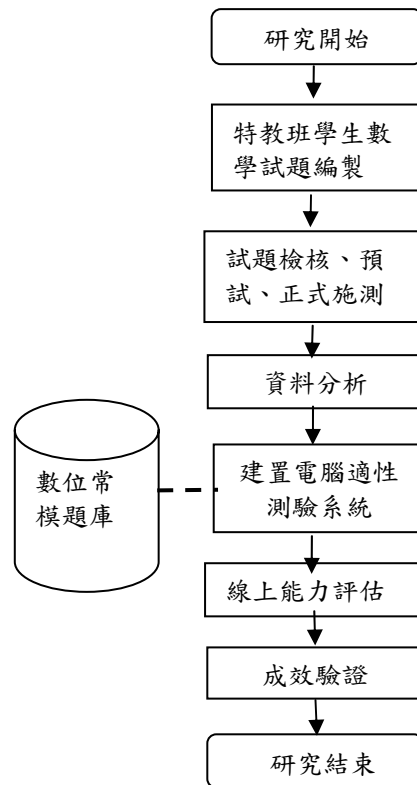


圖 3.1.1 網路適性測驗研究分析流程圖

3.2 研究對象

根據前節所述的研究方法問題，本研究擬採用的研究對象係為台灣地區高職特教班一、二年級學生之實用數學能力為研究對象。並以下列方式進行：

3.2.1 預試

本研究的預試樣本，以隨機選取高中職特教班一、二年級特教班學生為樣本，共計抽取 6 班 58 名學生為本研究的有效樣本。本預試工作，係在民國 94 年 1 月底之前，完成所有的書面施測，並根據施測結果由資深特教教師及國立台中師院教授共 5 人進行試題審核、命題內容、文句措辭、編排方式等提供意見，由研究者再加以修改。

3.2.2 正式施測

本研究的正式樣本，係以預試篩選試題後之正式試卷，於民國 94 年 2 月 15 日至 94 年 3 月 5 日之間，以校為單位隨機群體抽樣後全查方式向台灣地區 23 個縣市附設高職特教班之國私立高中職共 37 所學校共計 100 班發出 1277 份試卷，回收有效試卷樣本數 1108 份，回收率 86.77%。後續的資料分析工作，係以該有效樣本數作為分析的依據。

3.3 研究工具

本研究使用之研究工具有：測驗試卷、硬體設備、系統開發軟體工具等，茲將各種工具敘述如下：

3.3.1 測驗試卷

本研究使用實用數學科測驗試卷，其主要目的在於利用測驗施測結果，蒐集各試題參數，建立 ICC

曲線及訊息函數，以最大訊息量方法為選題核心機制，提供建置「線上題庫系統」。

3.3.2 硬體軟體設備

- (1). 硬體設備：Pentium4, 2.4GHz, 1GB RAM, Ultra320 SCSI HD 80GB, Intel Pro/100 網路卡, 傳輸效率 100Mbps。
- (2). 作業系統：Microsoft Windows Server 2000。Microsoft Windows Server 2000 是微軟較穩定的伺服器作業系統，不論是調配、管理和使用，都更為容易，且提供許多重要的新安全性功能和改善。
- (3). 資料庫：Microsoft SQL Server 2000。Access 2000 是微軟 Office 系列軟體中的一員，是一種控管非常嚴密的資料庫系統，適合於任何須要的資料庫。
- (4). Web Server：Microsoft Internet Information Services 6。為了增加網頁伺服器的安全性，IIS 6.0 所提供的網頁伺服器在可靠性、產能、連接性和整合性方面都是首屈一指的，其中包括容錯、要求佇列、應用程式狀態監視、自動應用程式回收、快取等等功能。
- (5). 網頁程式語言：Microsoft ASP.NET。ASP.NET 是一種以 Common Language Runtime 為基礎而建置成的程式設計架構，可用在伺服器上建置強大的 Web 應用程式。ASP.NET 提供幾項比過去 Web 開發模型還重要的優勢 - 它可讓您更輕鬆地撰寫程式碼，並提供增強的效能、安全性、強大功能和彈性。

3.3.3 測驗系統架構

本測驗系統的架構，包括六個模組：

- (1). 管理者模組：管理整個系統的教師使用。
- (2). 專家教師模組：進行試題的選取與組成試卷檔。題庫進行新增、刪除、修改、查詢及舉行施測等工作。
- (3). 審題教師模組：可以審查所出題目。
- (4). 命題教師模組：可以出題。
- (5). 任課教師模組：由系統管理者指派任課教師。
- (6). 學生模組：參與線上測驗評量和成績查詢。

3.4 實驗設計

本節的目的，旨在敘述將上述發展的鑑定模式置於網際網路上，供對個別受試學生的施測診斷之用，進而幫助教師或學生本身進行能力鑑定之分析。

3.4.1 網站架設

首先，先行架設網頁及網站，網址為：
<http://163.17.34.17/irt>

3.4.2 上網連結

接著，將前述發展的實用數學測驗的試題，登錄網頁題庫，使其能夠在網路上運算，幾經測試

後，才定稿完成建置工作。

3.4.3 線上測試

以 60~80 名高職特教班學生受試者上線接受測驗，收集其答題歷程產生之答題組態及能力值估計資料。

3.4.4 比較是否降低答題數

檢驗是否能減少答題數，在一定的能力估計誤差值下，仍能估算出受試學生能力或預測或解釋試題上的表現情形。

本研究測驗的中止的題數是由預定的前後測量標準誤差距(standard error difference of estimation)來決定。其計算方法可以公式(3.1)表示：

$$SED(\hat{\theta}_{i+1}) = SE(\hat{\theta}_i) - SE(\hat{\theta}_{i+1}) \quad (3.1)$$
$$i = 1, 2, \dots, n-1$$

$\hat{\theta}_i, \hat{\theta}_{i+1}$: 分別指前後答題所測得的估計能力值

n : 指最大施測試題數

本研究依實際題庫資料來決定估計標準誤差距來設定題數中止標準，並以「選題施測過程已答題 10 題以上」，且所得受試者的前後題能力估計標準誤之差距達設定「閾值 ≤ 0.002 」所作答的題數能力，來做統計驗證。設定需答題 10 題以上係為考量試題周延性及內容效度，且避免估計能力值變動在未收斂前，恰巧落在閾值內產生誤判情形。

4. 結果與討論

以下將就本研究線上測驗，分一、測驗系統操作說明及二、適性測驗後測結果兩部份說明結果。

4.2.1 測驗系統

本研究利用所開發的「線上題庫系統」，由於結合了網際網路，建置在全球資訊網的平台上，可提供不同管理權限者，審題、命題、修改試題，並進行試題管理，達到組卷的功能。

受試者隨時上線作答，不受時空的限制，透過網際網路的連結，至少可以讓不同地區多位答題者同時上線作答，除初始題目採隨機選題外，採用估計能力值最大訊息選題法作為核心選題方式，因此每位受試學生其答題組別及順序都不同，完全依據學生每題答完後計算之能力，選擇適性題目，作答完成後，電腦會產生歷程紀錄提供給受試者。

在學生測驗結果，我們記錄學生答題歷程，包括答題次序、答對率及更重要的是將利用電腦強大運算功能，即時根據線上受試者當時進行之答題組型，所產生估計能力值，並根據試題參數，選擇下一個訊息量最大題目，其與一般題庫大量建立試題隨機選題方式不同，係為智慧型選題，故稱為適性測驗。

4.2.2 測驗分析

最後適性測驗另一項重要功能，即運用選題機制確能在固定答題或減少施測題數下，達到預測

受試者能力特質之效果，並能節省施測時間。

為證明本適性測驗系統建置成效，於 94 年 6 月 20、21 日兩天以國立霧峰農工、台中高農、新竹高工等 3 校共 6 班高職特教班學生進行遠距線上適性測驗，其中 68 位有效受試者之答題及歷程，經系統即時得到結果分析，以下將就代號 p20 之受試同學選題及能力值估計歷程說明。

由下表 4.2.1 可知，p20 同學答完全部 26 試題所得

到之能力值 $\hat{\theta}_{26}^{(26)}$ 為 0.44，由表內資料可知答題順序 23 其估計標準誤之差為 0.0018 符合達到「閾值 ≤ 0.002 」停止施測條件，所得之作答順序第

23 題估計能力值 $\hat{\theta}_{26}^{(23)} = 0.48$ ，答題順序 24、25、26 則穩定均為 0.44。

表 4.2.1 電腦適性測驗答題能力估算歷程表

答題順序	答題題號	答題結果	能力值 θ	$t(\theta)$	估計標準誤	標準誤前後差	訊息函數	正確答案	學生答案
1	7	正確	3	0.98	62.0410	-	0.0003	3	3
2	2	錯誤	-0.24	0.38	3.5197	58.5213	0.0807	3	1
3	20	正確	0.36	0.86	0.8509	2.6688	1.3810	2	2
4	3	正確	0.6	0.79	0.5288	0.3221	3.5756	1	1
5	16	正確	0.76	0.54	0.4307	0.0981	5.3903	3	3
6	17	錯誤	0.56	0.51	0.3960	0.0347	6.3755	2	3
7	28	正確	0.6	0.23	0.3953	0.0007	6.3982	1	1
8	24	正確	0.68	0.69	0.3713	0.0241	7.2552	2	2
9	13	錯誤	0.56	0.5	0.3520	0.0193	8.0705	3	2
10	15	正確	0.6	0.82	0.3351	0.0169	8.9046	3	3
11	26	錯誤	0.48	0.57	0.3234	0.0117	9.5595	1	2
12	4	正確	0.52	0.91	0.3129	0.0105	10.2133	4	4
13	19	正確	0.52	0.82	0.3052	0.0077	10.7372	2	2
14	8	錯誤	0.36	0.91	0.2964	0.0087	11.3794	3	2
15	11	正確	0.36	0.9	0.2891	0.0073	11.9644	1	1
16	22	正確	0.4	0.9	0.2836	0.0055	12.4295	1	1
17	9	正確	0.44	0.53	0.2789	0.0047	12.8557	4	4
18	5	錯誤	0.4	0.56	0.2749	0.0040	13.2311	2	4
19	14	正確	0.4	0.97	0.2707	0.0042	13.6426	2	2
20	25	正確	0.4	0.94	0.2670	0.0037	14.0245	4	4
21	21	正確	0.44	0.86	0.2641	0.0029	14.3328	4	4
22	6	正確	0.44	0.47	0.2612	0.0029	14.6561	1	1
23	23	正確	0.48	0.97	0.2595	0.0018	14.8552	4	4
24	10	錯誤	0.44	0.25	0.2585	0.0009	14.9634	3	2
25	1	正確	0.44	0.89	0.2576	0.0009	15.0734	3	3
26	12	錯誤	0.44	0.38	0.2568	0.0007	15.1584	2	1

由整體 68 名高職特教班學生統計資料來分析，在平均作答部分，總題數 26 題，適性測驗符合閾值終止施測題數之平均作答題數為 19.85，亦即平均節省施測題目 6.15 題。估計能力值部份，68 名受試者 26 題全答估算之平均能力值為 -0.37，適性達終止條件之平均能力值為 -0.35，計算兩者相關係數達 0.9613 強。測驗終止的能力估計均方根差 $RMSD(\theta)$ (root measurement of square difference)，以公式(4.1)表示如下：

$$RMSD(\theta) = \sqrt{\frac{\sum_{k=1}^n (\hat{\theta}_{ks} - \hat{\theta}_{kn})^2}{n}} \quad (4.1)$$

其中 $\hat{\theta}_{ks}$ ：指依據第 k 位受試者適性終止施測題數所得估計能力值

$\hat{\theta}_{kn}$ ：指第 k 位受試者答完全部測驗

題目所得到之能力值

n：指受試總人數

本研究依實際題庫資料來決定估計標準誤差距來設定題數中止標準，並以「選題施測過程已答題 10 題以上」，且所得受試者的前後題能力估計標準誤之差距達設定「閾值 ≤ 0.002 」所作答的題數能力，來做統計驗證。設定需答題 10 題以上係為考量試題周延性及內容效度，且避免估計能力值變動在未收斂前，恰巧落在閾值內產生誤判情形。

另本測驗 $RMSD(\theta)$ 為 0.214257，此數據比較陳俊宏(民 93)所分析得到 $RMSD(\theta)$ 約在 0.3 至 0.4 之間，誤差更小，本研究符合其所稱優質題庫一致性的要求。

另外以作答時間來看，26 題全部作答時間平均為 18.88 分，若平均可節省施測試題 6.15，也可以換算為平均節省施測時間 4.47 分鐘，亦即節省 23.65% 的受試時間，因此電腦適性測驗確可達到節省作答時間目的。

5. 結論與建議

5.1 結論

本研究 CAT 系統經由對高職特教班學生於網際網路線上後測分析結果，確實能依據個別學生能力不同，由系統適性選題，能打破固定出題順序並節省傳統紙筆作答試題數，達到節省題試及受測時間，又能有效估算出受試學生能力特質。

本節歸納提出本研究的結論如下：

- (1). 電腦化適性測驗符合特殊教育「個別化」精神：本研究適性測驗證明能克服目前特殊教育學習目標太多無法全部施測情形，而改採以特殊教育「個別化」測驗方式，每個學生線上施測的題目順序均不同，以最符合個別學生能力估計且能得到最大訊息量之試題進行評量。
- (2). 電腦化適性測驗符合特殊教育「適性化」精神：本研究適性測驗結果，節省施測試題及作答時間平均達 23.56%，有效控制能力估計均方根差，能準確評估個別智能障礙類學生數學能力，瞭解學生起點能力作為教學參考。

5.2 建議

茲針對本研究結果的研發心得與後續的追蹤研究，提出以下建議：

- (1). 本研究需使用電腦上網測試，因此僅適用於輕度智能障礙學生，對於伴隨其他障礙或多重障礙類學生，施測教師應視應試需求，提供輔助性服務，如試題提示、報讀、協助操作等，否則電腦操作因其功能缺損障礙，而成為能力評估的門檻。
- (2). 目前特殊教育所謂 IEP 皆強調學習輔導資料記載或學習目標設定，而兩者之間關連性卻缺乏所謂診斷評量轉換機制，運用 IRT 所謂知識

結構理論來分析上下位概念構圖，診斷出對學生的迷思概念加以補救教學，此部份仍待後續有興趣研究者深入探討。

參考文獻

- 王崇懋(民 94)。網路資源應用於特殊教育教學現況分析。南華大學社會學研究所網路社會學通訊期刊，第四十一期。
- 王華沛(民 91)。輔助科技之應用。台北：中華民國輔助科技促進職業重建協會。
- 王寶墉(民 84)。現代化測驗理論。台北：心理出版社。
- 朱經明(民 86)。特殊教育與電腦科技。台北市：五南。頁 56-65，251。
- 余民寧(民 86)。試題反應理論之介紹—測驗理論之發展趨勢(一)。研習資訊，第 8 卷第 6 期，頁 13-15。
- 余民寧(民 86)。試題反應理論的介紹—能力與試題參數的估計。研習資訊，9(3)，頁 6-12。
- 余民寧(民 86)。試題反應理論的介紹—訊息函數。研習資訊，9(6)，頁 5-9。
- 余民寧(民 86)。試題反應理論的介紹—電腦化適性測驗。研習資訊，10(5)，頁 5-9。
- 李翠玲(民 88)。IEP 的理念與其問題。載於竹師特教簡訊，28，1。
- 林寶貴主編(民 89)。特殊教育理論與實務。台北：心理出版社。
- 洪榮昭(民 88)。電腦輔助教學之設計原理與應用。台北：師大書苑。
- 陳俊宏(民 93)。電腦適性測驗之模擬研究。國立台中師範學院教育測驗統計研究所理學碩士論文，頁 20-24。
- 陳新豐(民 88)。多媒體線上適性測驗系統發展及其相關研究。國立臺南師範學院國民教育研究所碩士論文。
- 陳麗如(民 87)。電腦化適性測驗之題庫品質管理策略。國立臺灣師範大學資訊教育研究所，碩士論文。
- Birnbaum, A. (1968). *Some latent trait models and their use in inferring an examinee's ability*. In F. M. Lord & M. R. Novick, *Statistical theories of mental test scores* (chapters 17-20). Reading, MA: Addison-Wesley.
- Hambleton, R. K. (1985). *Item response theory: Principles and application*. Boston: Kluwer Nijhoff.
- Lord, F. M. (1977). *Practical applications of item characteristic curve theory*. *Journal of Educational Measurement*, 14, 117-138.
- Reckase, M. D. (1973). *An interactive computer program for tailored testing based on the one-parameter logistic model*. Paper presented to the National.
- Swaminathan, H. (1983). *Parameter estimation in item response models*. In R. K. Hambleton (ed.). *Application of item response theory*. Vancouver: Educational Research Institute of British Columbia.
- Urry, V. W. A. (1970). *Monte Carlo investigation of logistic test models*. WestLafayette, IN: Unpublished doctoral dissertation, Purdue University.
- Weiss, D. J. (1974). *Strategies of adaptive ability measurement* (Research Report 74-5). University of Minnesota, Department of Psychology, Psychometric Methods Program.