導入布魯姆認知分類及量尺定錨法之適性化評量系統 以 C 語言為例

張儀興　　陳煥文
南台科技大學資訊管理所
yhchang@mail.stut.edu.tw　　M9590207@webmail.stut.edu.tw

摘要

本論文導入布魯姆認知分類修正版及量尺定錨法為基礎提出一適性化評量系統，AAS，其主要設計概念是利用布魯姆認知分類修正版與量尺定錨法進行試題質與量的分析，以得到一個試題精確的特徵資訊，並結合學習者的測驗歷程檔來達到適性化的評量。此機制主要的特色為提供一個電腦化的線上題庫使教學者方便管理編輯測驗試題，試題知識庫內所記錄的資訊也能供學習者針對某一課程觀念進行學習，並能藉由經分析之試題來達成客觀、精確的評量。

關鍵詞：適性化評量系統、布魯姆認知分類、量尺定錨法、試題知識庫

Abstract

An adaptive assessment system, AAS, based on the revised Bloom's taxonomy and scale anchoring is proposed. The revised Bloom's taxonomy and scale anchoring is used to analyze the quality and quantity of items and get the items' characteristic information. Finally, the items' characteristics and learners' test portfolio are used to achieve the adaptive assessment. The features of the AAS are: providing a on-line item base to help the teachers manage conveniently items; giving the learners to learner some specific curricula again according to the information stored in the item base; ,and achieving the objective, precise assessment by analyzing the items' characteristics.

Keywords: Adaptive assessment system, Scale anchoring, Bloom's taxonomy, Item knowledge base

1. 前言

數位學習環境具有不受時間與空間限制的優越特性，因此已受到廣泛的重視，尤其在寬頻普及的時代裡，許多企業組織用來作為員工教育訓練以提昇其競爭力。其中，不論在傳統或電腦化的教學，測驗與評量一直在教學過程中有相當程度的重要性。設計良好的試題可準確的反應出學習者學習的程度，以供教學者掌握學生的學習狀況，進而調整教學的策略。然而其評量出的學習情況往往是不夠客觀精確的，導致此現象的原因是由於對於測驗試題的設計，皆沒有考量到與其教學目標是否一致；另外在試題困難度的訂定上，都是依照專家教學者由其多年的教學經驗主觀的分級，這樣的方式可能無法精確的反映出學習者真正的學習情況。

基於上述理由得知，要提供學習者一個適性化的評量系統，是否能準確的評量出學習者的學習成效將學習者習得的知識成果或技術能力呈現出來，測驗試題扮演著關鍵性的角色，因此本研究中應用量尺定錨法將試題的難易度對應至總分的量尺上，更明確的顯示層級的意義；再以布魯姆認知分類修正版理論為基礎，更具體來描述試題的測驗功能。最後以經過分析後之試題並結合測驗歷程檔以建構一個適性化評量系統。

2. 文獻探討

2.1 量尺定錨法

量尺定錨是在不同的量尺點上，將受試者表現的特徵展示出來，其形式是選擇一些量尺分數作為定錨點，並以統計方式選擇試題作為該點的定錨試題。定錨試題是指達該定錨點的受試者，有較高機會答對的試題，而低於該點的受試者則不然。Beaton & Allen[11]針對量尺定錨提出兩種計算方式-直接法 (the direct method) 與平滑法 (the smoothing method)。兩種方法皆需先取得受試者的作答反應與成績。直接法是採用不同定錨層受試者的答對率，作為試題歸屬時的決定標準；平滑法則是應用試題反應理論(IRT)，建立試題特徵曲線，再計算各試題在不同定錨點的答對率。兩種方法相異之處在於求答對率的程序上。雖然直接法計算簡單，但考量樣本數不多情況時，答對率的計算會產生較大的誤差[6]，因此本研究主要參考量尺定錨的平滑法。

2.2 試題反應理論

試題反應理論(IRT)是以數學式表示受試者能力與試題參數間的關係，而這個數學關係即是大家所熟悉的機率，以受試者的答對機率為縱軸，受試者的能力為橫軸所畫出來的曲線即為試題特徵曲線(item characteristic curve，簡稱 ICC)。依照試題

特徵曲線內所含參數個數，一般可將 IRT 模式分為單參數對數模式(1-prameter logistic model，簡稱 1PLM)、雙參數對數模式(2PLM)和三參數對數模式(3PLM)，如表 1 所示，其中 Pi(θ)代表能力值為θ之受試者答對第 i 題的機率，D 為常數 1.7，θ為受試者能力值，i 為試題數[2]。

表 1 試題反應理論的主要模式對照表

IRT模式	公式	參數
單參數對數模式(1PLM)	$P_i(\theta)=\frac{\exp D(\theta-b_i)}{1+\exp D(\theta-b_i)}$	b_i：難度
雙參數對數模式(2PLM)	$P_i(\theta)=\frac{\exp Da_i(\theta-b_i)}{1+\exp Da_i(\theta-b_i)}$	a_i：鑑別度、b_i：難度
三參數對數模式(3PLM)	$P_i(\theta)=c_i+(1-c_i)\frac{\exp D(\theta-b_i)}{1+\exp D(\theta-b_i)}$	a_i：鑑別度、b_i：難度、c_i：猜測度

習慣上常將單參數模式又稱 Rasch 模式。經王文中[1]研究指出，只有 Rasch 模式才是測量的模式，而二參數模式和三參數模式只有統計模式，沒有測量的意義，亦即 Rasch 模式所計算出的試題難度值及受試者能力值，具有等距(interval)和客觀(objectivity)的特性，因此可以相互比較。基於此，本研究將採用 Rasch 模式來分析作答結果，以估計試題的難度、受試者能力值與答對機率等特徵資料。

2.3 測驗歷程檔

歷程檔案(Portfolio)可以意指任何東西，從累積性的學生寫作文件到私人的剪貼簿都屬之，最早被廣泛應用在藝術上，有效的記錄學習者的作品集[12]。學習歷程記錄，其應用已成為近年來教育上新興的發展趨勢之一，它以特定的目的或用途來蒐集學習者學習過程中各面向的資料，以反應或顯現學習者的真實表現及進步或改變的情形[5]。

2.4 布魯姆認知分類理論

美國教育學者 Bloom 等人在 1956 年提出，將教學目標分成認知、情意、技能三大領域。此分類系統的主要特色，它是一種由具體到抽象，由簡單到複雜且具有內在階層性與累積性，因此適合作為衡量教材內容、教學活動或評量內容之均衡性指標。而此認知分類架構經過多年的使用，許多學者提出需要修正檢討的呼籲，如：使用行為目標對教學者造成限制、忽略了無法明確陳述的重要目標、視個別目標為獨立的目的欠缺關聯性、類別的區分性不足易造成混淆。加上近年來對於學習課題，強調學習者「知」(Knowledge)和「如何思考」(how they think)的概念[8]，終於在 2001 年由 Anderson 和 Krathwohl 等人提出了布魯姆認知分類修訂版[10]。圖 1 為新舊版本改變情形其中知識向度的部份代表了學習者學習成果的累積，而認知歷程向度的部份則代表了學習者在學習時思考的方式與過程。

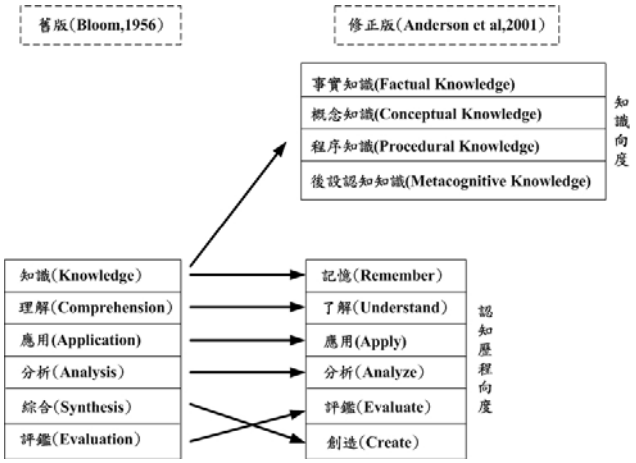

圖 1 布魯姆認知分類新舊版本改變情形[10]

3. 適性化評量系統

本適性化評量系統（Adaptive Assessment System，AAS）其系統架構圖如圖 2 所示，使用者透過網頁介面使用本系統。系統中包含了學習者資訊模組、試題知識庫模組、評量模組及資料庫分別介紹如下：

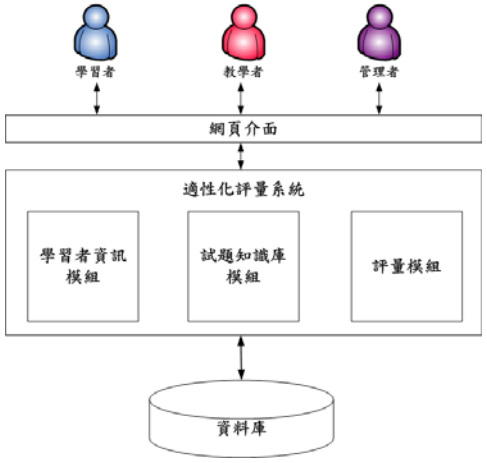

圖 2 適性化評量系統架構圖

3.1 學習者資訊模組

學習者資訊模組運作流程如圖 3 所示，此一模組主要功能在記錄學習者的測驗歷程以供評量模組提供學習者適性化評量之依據與依據所評量出學習者的學習情況給予學習者適性化的學習建議；此模組並且提供教學者關於學習者之作答反應之統計性資訊做為調整教學方針之參考。學習者資訊模組提供之適性化學習建議範例如下：

「你在結構體這個課程很多地方學習並不夠熟練

哦！尤其是在應用與創造方面的能力需要加強。」

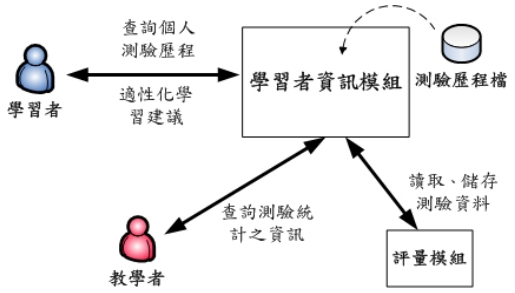

圖 3 學習者資訊模組運作流程

3.2 試題知識庫模組

試題知識庫模組主要功能為提供評量模組所需之測驗試題，試題需經過「質」與「量」的分析步驟後存放於試題知識庫中，並加上該試題解題知識儲存於試題知識庫中提供學習者針對其不足處查詢該試題之解題相關知識，提供二次學習的機會。另外教學者可查詢與修改試題知識庫內的試題相關資料。教學者依據測驗歷程檔中所記錄之學習者的作答反應的執行試題量的分析即試題執行量尺定錨法，得到試題難易度的「量化指標」後。提供給評量模組做更精確的適性測驗幫助教學者有效的掌控學習者的學習情況，以做為調整教學內容或針對個別學習者的弱點進行補救教學之根基。

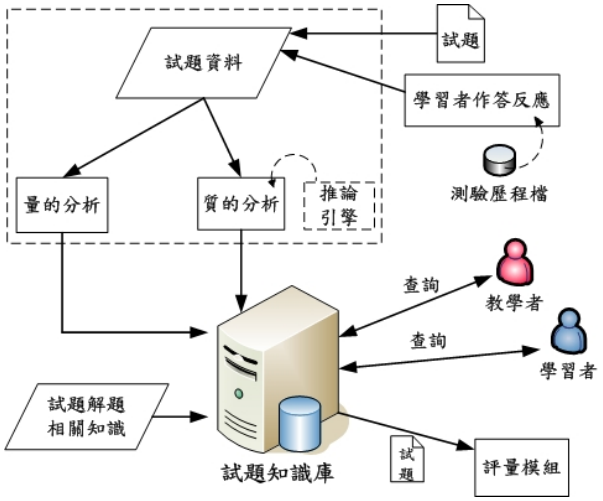

圖 4 試題知識庫模組流程

3.2.1 推論引擎

在試題「質的分析部分」將以布魯姆認知分類修正版做試題分析，參考鄭蕙如、林世華[9]研究中將九年一貫課程數學領域分段能力指標以動詞、名詞兩部分加以分解，動詞的部分套在認知歷程向度，名詞則套在知識向度。另外在沈孟祥[3]的研究中動詞的部分套在認知歷程向度，以問句參考句型套在知識向度。本研究參考過去學者之研究設計一推論引擎如圖 4 小虛線方格標示處，此推論引擎主要功能為自動化的分類出試題的布魯姆認知分類的認知歷程與知識二個向度，以輔助教學者執行試題「質的分析部分」減輕教學者的工作負荷，此推論引擎運作流程如圖 5 所示，將試題題目拆解成代表認知歷程向度的「動詞」詞彙與代表知識向度的「問句關鍵字」二部份。接著依此質化指標儲存於試題知識庫中，教學者可依需求再檢視修改。

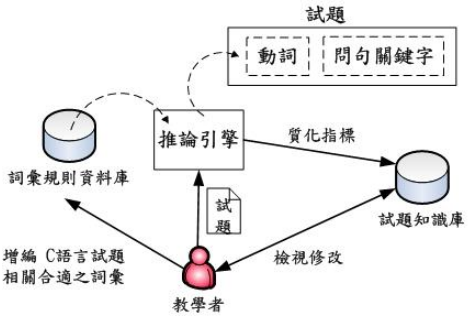

圖 5 推論引擎運作流程

3.2.2 以布魯姆認知分類修正版做試題質的分析

由布魯姆認知分類修正版的知識向度與認知歷程向度，可形成二向度次類別分類表如表 2 所示，由於布魯姆認知分類具有檢視教學目標、剔除設計不良試題之功能，因此相當適合作為分析試題之用；且分析後所得之認知歷程向度、知識向度分類指標，也可提供給學習者讓學習者了解本身所缺乏之能力。

表 2 二向度次類別分類表

知識向度		認知歷程向度																				
主類別	主類別	1記憶		2了解							3應用		4分析			5評鑑		6創造		
	次類別	1 再認	2 回憶	1 詮釋	2 舉例	3 分類	4 總結	5 推論	6 比較	7 解釋	1 執行	2 實行	1 辨別	2 組織	3 歸因	1 檢查	2 評論	1 產生	2 計畫	3 製作
A 事實	A 術語																					
	B 特定細節和元素																					
B 概念	A 分類和類別																					
	B 原則和通則																					
	C 理論 /模式/結構																					
C 程序	A 特定學科的技能和演算																					
	B 特定學科的技術和方法																					
	C 運用規準																					
D 後設認知	A 策略																					
	B 認知任務																					
	C 自我																					

試題質的分析演算法如下：

Input：試題題目。

Output:試題質化指標。

步驟 1： 教學者針對其教學目標設計試題。

步驟 2： 教學者以系統提供之推論引擎輔助執行布魯姆認知分類修正版做試題質的分析部分。

步驟 3： 得到試題的質化指標（認知歷程向度、知識向度的分類指標）；於此

步驟教學者即可發現並剔除設計不良與其教學目標不符的試題。

步驟 4： 最後教學者針對該試題加入欲提供給學習者的試題解題相關知識。

本研究將以 C 語言為主要教學課程，以下部份舉例以動詞與問句關鍵字字彙判斷認知歷程與知識向度之試題：

（一）認知歷程向度--**記憶**，知識向度--**事實**

題目：請指出 C 語言翻譯成機器語言的方式為?

動詞：指出（1.1 再認）

問句關鍵字：的方式為(Ab.特定細節和元素的知識)

（二）認知歷程向度--**了解**，知識向度--**概念**

題目：若要宣告一變數，以下何者不能用作使用者識別字？

動詞：不能用（2.6 比較）

問句關鍵字：變數、使用者識別字（Bb.原則和通則的知識）

（三）認知歷程向度--**應用**，知識向度--**程序、概念**

題目：請依據下列程式可修改合併成一條 if 判斷式為?

```
if(a<=b) x = 123;
else y = 456;
if(a>b) printf("@@@@ y = %d \n",y);
else printf("$$$$ x = %d \n",x);
```

動詞：修改（3.1 執行）

問句關鍵字：合併成（Ca.特定學科的技能和演算的知識）、if 判斷式（Bb.原則和通則的知識）

（四）認知歷程向度--**分析**，知識向度--**概念**

題目：請依據以下程式，分析下列程式屬於何種流程控制?

```
void main(void)
{   char a;   prfintf("你尚未登入…");
do{   printf("是否重新登入?(y/n) ") ;
a=getchar() ;   getchar() ;   }
while((a!= 'y')&&(a!= 'n'));   }
```

動詞：分析（4.1 辨別）

問句關鍵字：何種流程控制（Ba.分類和類別的知識）

（五）認知歷程向度--**評鑑**，知識向度--**程序**

題目：以下 C 語言程式中多處有錯，請判斷程式中錯誤部分及原因。

```
main;   {
float a,b,c,d;
a=7.0;b=1.3;c=9.0
d=a*b*c;
printf(“%f\n”,d)   }
```

動詞：判斷（5.1 檢查）

問句關鍵字：錯誤部分及原因（Cb.特定學科的技術和方法的知識）

（六）認知歷程向度--**創造**，知識向度--**後設認知**

題目：請設計撰寫一程式，功能為要求使用者輸入一整數給一整數資料型態變數，數值範圍 0～32767 若不滿足此條件，則要求使用者重新輸入，並從個位數開始輸出此一整數變數中每一位上的數字每輸出一數字即換行。

動詞：設計（6.2 計畫）

問句關鍵字：撰寫一程式（Da.策略知識）

3.2.3 以量尺定錨法做試題量的分析

參考過去傳統數位學習系統和我們過去的研究中，發現到在測驗試題的難易度等級區分，大多以教學者的經驗主觀的訂定。考慮到這樣的方式可能無法很精確依照學習者的學習情況提供適合的評量，所以藉由對試題做量的分析，以得到客觀、精確訂定之試難易度指標。

試題量的分析演算法如下：

Input：學習者作答反應資料。

Output：試題量化指標。

步驟 1： 從測驗歷程檔中取得作答反應資料，並整理成 Rasch 模式分析時所需符合的格式。

步驟 2： 執行 Rasch 模式的分析，擬合出一條試題特徵曲線。

步驟 3： 依據試題特徵曲線，估計試題在各定錨點的答對率。

步驟 4： 決定符合最低定錨層的試題，其答對率的規準（本研究在定錨試題答對機率門檻標準主要參考[4]設為 65%）。

步驟 5： 決定符合其它定錨層的試題，其答對率的規準。

步驟 6： 對各定錨層，依據解決該試題所需的知識概念與程序技巧做歸納。

本研究以 θ_{10}、θ_{20}、θ_{30}、θ_{40} 分別代表四個定錨點分數，Over40 則代表試題在定錨點 40 的試題答對率仍本達門檻標準。為了表達試題的難易度，本研究按照定錨點層級由低由高，將試題劃分成簡易、偏易、中等、偏難、困難五個等級。

3.3 評量模組

在介紹評量模組前，首先說明本研究提供學習者之試題題型。我們發現到一般數位學習平台，通常都只提供選擇題形式的測驗試題，這樣的方式可能存在著一些評量上的盲點，例如使用選擇題的方式出題，很難以評量出學習者關於創造等較高層次之能力。參考學者葉連祺研究指出每種類型之測驗試題適合評量的認知目標層次並不相同，選擇題型

較適合評量學習者知識、理解、應用和分析等層次，論述題型則適合評量綜合、評鑑和創造等能力[7]。本研究採用選擇與非選擇題的出題方式，來輔助關於選擇題在評量上的盲點。另外，由於目前電腦化測驗與評量的方式，尚無法對於問答題型進行自動批改，故評量流程較選擇題型有所不同流程如圖 6 所示：

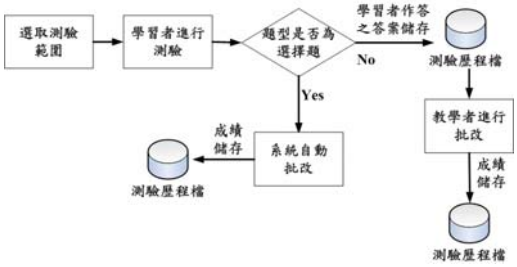

圖 6 不同題型之評量流程

在此評量模組是以試題知識庫裡，經過質與量分析程序的試題來做一個適性化的評量流程如圖 7 所示。由於儲存於試題知識庫裡的測驗試題，皆經過量尺定錨法分析試題而得到一個精確的試題分層級的「量化指標」，此一「量化指標」代表著該測驗試題的難易度，而經由布魯姆認知分類修正版分析後所得到該測驗試題之「質化指標」代表學習者所學習到的能力程度。接著即可對學習者做一個前測評量，並將測驗結果儲存於學習者資訊模組裡的測驗歷程檔，往後即可依學習者的程度、能力提供適性化的評量。

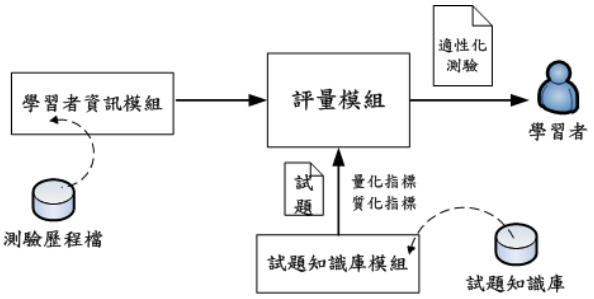

圖 7 評量模組流程

以下說明評量模組提供適性化評量之選題策略：

步驟 1:

針對課程單元對學習者做一個前測的評量流程，以得知該學習者目前大概的學習情況如圖 8 所示。前測的題目等級是挑選在一個課程評量範圍裡中等難度（θ30）並依據試題布魯姆認知分類認知歷程向度記憶、了解、應用、分析、評鑑、創造六個類別分別各挑選一題給該學習者做一個前測評量，在往後各階段之評量也都依據認知歷程向度六個類別各挑選一題，若某一類別並無試題則略過此一類別之試題。

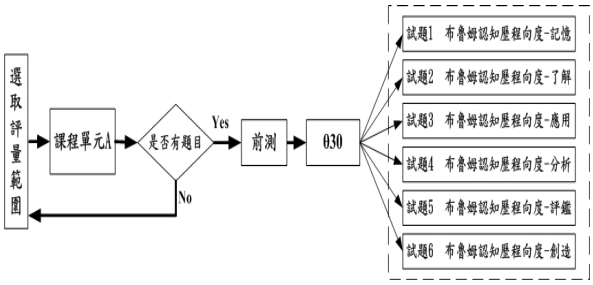

圖 8 選題流程步驟 1

步驟 2:

如圖 6 所示，在學習者前測作答結束後，系統依據學習者在前測的作答反應來決定提供給學習者評量的題目難度是否上升一個等級或下降一個等級。例如該學習者共作答 6 題 4 題答對，2 題答錯則有 67%的試題為正確的作答反應，依據圖 9 所示即提升一個等級難度偏難（θ40）之試題給學習者，反之若試題為正確的作答反應低於 65%則下降一個難度等級提供偏易（θ20）之試題給學習者。往後學習者在此一課程 A 裡的評量流程如前述之方法，依學習者的作答情況動態的調整提供給學習者的試題困難度，此一課程評量的結束條件為學習者已做完在此選取的課程單元範圍內最高或最低等級之試題，另外在圖 9 中以虛線特別標明之定錨點θ40 在本研究將此一等級之試題定為一個判斷學習者是否到達精熟的學習狀態的定錨點，例如某一學習者在偏難（θ40）等級之試題的正確作答反應高於 65%，則系統即認定該學習者對於此一課程單元範圍的學習情況已到達精熟的學習狀態。

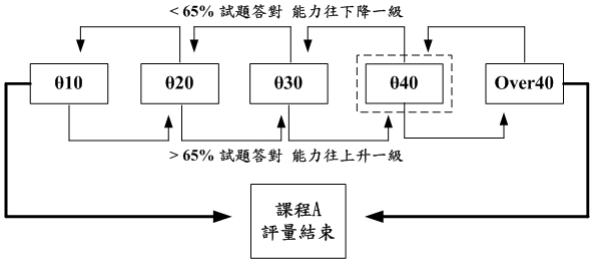

圖 9 選題流程步驟 2

3.4 資料庫設計

3.4.1 測驗歷程檔資料庫

學習者帳號：識別學習者之用。

測驗題目：學習者已做答之題目。

測驗分數記錄：此分數記錄為二元方式計分，記錄學習者對於該試題的反應；作答正確為 1 反正為 0。

學習者作答內容：記錄學習者作答內容，以供批改。

學習者能力等級記錄：記錄學習者對於此一課程範圍的精熟程度。

3.4.2 試題知識庫

試題編號:每道試題編碼方式均依"課程主題"+"流水號"所編碼。

課程主題：試題所屬之 C 語言課程單元。

題型：為選擇題或其它題型如問答題填充題等等。

試題內容：為試題的主要敘述部分，其可能包含文字、圖示、表格等。

量尺定錨指標：以「量尺層級來表示」試題經過分析後的量化指標。

布魯姆認知分類指標：顯示解決問題的能力類別，以布魯姆分析得到的結果，並以文字方式精確的描述試題涵蓋的測驗功能之質化指標。

試題解題相關知識：提供學習者關於該學習主題額外的參考知識、解題觀念詳解等等。

3.4.3 詞彙規則資料庫

認知歷程向度動詞：認知歷程向度次類別分類，C 語言試題相關動詞詞彙。

知識向度問句關鍵字：知識向度次類別分類，C 語言試題相關問句關鍵字詞彙。

表 3　認知歷程向度動詞範例

主類別	次類別	動詞
記憶	再認(1.1)	指出、識別出、指名・・・

表 4　知識向度問句關鍵字範例

主類別	次類別	問句關鍵字
概念	原則和通則(Bb.)	變數、ｉｆ判斷式、陣列・・・

4. 結論

電腦化評量一直是數位學習領域中重要的一個議題。本篇研究主要結合量尺定錨法與布魯姆認知分類修正版理論，針對試題加以分級及試題內容效度的剖析，並結合測驗歷程檔設計一套適性化評量系統，依教學者所設定之目標學習者可藉由，試題知識庫、測驗歷程檔自我檢視學習成果來完成學習目標。綜合以上所述本論文結合量尺定錨法與布魯姆認知分類修正版理論於電腦化評量系統主要優點如下：

對試題進行質的分析，由於導入布魯姆認知分類修正版理論，能更明確描述試題的測驗功能。

以量尺定錨中的平滑法來進行量的分析對試題作更精確的層級歸屬。

本研究參考過去之研究設計一推論引擎期能有效輔助教學者執行試題質的分析，降低教學者的工作負荷。

本研究依據量尺定錨法與布魯姆認知分類理論分析得到之試題的「質化指標」與「量化指標」提出一「適性化的評量策略」來測驗學習者，更可以較準確的評估學習者程度，據此做為適性化學習的依據。

5. 誌謝

本研究由國科會計畫 96-2520-S-218-001-補助支持，特此誌謝。

參考文獻

[1] 王文中，"Rasch 測量理論與其在教育和心理之應用"，教育與心理研究，第二十七卷第四期，pp. 637-694，2004。

[2] 王寶墉，現代測驗理論，心理出版社，臺北市，1995。

[3] 沈孟祥,應用布魯姆教育目標認知分類理論設計發問工具之研究,佛光人文社會學院教育資訊學研究所碩士論文，未出版，2005。

[4] 洪瑞鎂,從「第三次國際數學與科學教育成就研究後續調查」探究台灣國二學生的數學基本能力,國立台灣師範大學數學研究所碩士論文,未出版，2001。

[5] 張基成、童宜慧，"網路化學習歷程檔案系統之建構與評鑑－一個電子化的真實性學習評量工具"，遠距教育，第十三、十四期，pp. 78-90，2000。

[6] 張儀興、許世豐，"設計及完成以量尺定錨法為基礎之試題知識庫系統"，TANET 2006 台灣網際網路研討會，十月，2006。

[7] 葉連祺，"教師自編紙筆式測驗試題類型之探討"，研習資訊，第十七卷第四期，pp. 42-53，2000。

[8] 葉連祺、林淑萍，"布魯姆認知領域教育目標分類修訂版之探討"，教育研究月刊，第一百零五期，pp. 94-106，2003。

[9] 鄭蕙如、林世華，"Bloom 認知領域教育目標分類修訂版理論與實務之探討 - 以九年一貫課程數學領域分段能力指標為例 "，台東大學教育學報，第十五卷第二期，pp. 247-274，2004。

[10] Anderson W. & Krathwohl, D. R.(Eds.), A taxonomy for learning, teaching, and Assessing：a revision of Blooms' Educational Objectives, New York, NY：Longman, 2001.

[11] Beaton, A. E., & Allen, N. L., "Interpreting scales through scale anchoring," Journal of Educational Statistics, Vol. 17, No. 2, pp. 191-204, 1992.

[12] Graves,D & Sunstein,B.S.(EDs.), Portfolio portraits, Portsmouth. NH:Heinemann, 1992.