

以 DINA 模式為基礎建置電腦化適性動態評量教學系統

王曉瑜	劉宇文	郭伯臣	鄭俊彥
國立臺中教育大學 教育測驗統計研究所 碩士生 sandy2910@yahoo.com m.tw	國立臺中教育大學 教育測驗統計研究所 碩士生 g4j4@yahoo.com.tw	國立臺中教育大學 教育測驗統計研究所 教授兼教育學院院長 kbc@mail.ntcu.edu.tw	國立臺中教育大學 教育測驗統計研究所 博士生 x7622@hotmail.com

摘要

傳統紙筆測驗只能知道試題對與錯，卻無法了解學生所欠缺的認知概念，因此認知診斷便是彌補傳統測驗之不足，但因學生人數多，導致學生在作答有誤時無法立即給予提示或回饋，倘若認知診斷能結合電腦適性化動態評量將能省時且省力。本研究是模擬平均每題 1.2、1.8、2.4、3.6 個概念屬性之 Q 矩陣以及 52 筆實徵資料之 Q 矩陣在 0.7、0.75、0.8、0.85、0.9 等五種終止條件下，其後平均屬性辨識率與省題率之結果，進而建置教學系統。由研究結果得知：模擬不同 Q 矩陣與實徵資料其終止條件對於平均屬性辨識率 (mean of each attribute accuracy) 作為評估指標，影響很大，且實徵資料在終止條件為 0.7 時，省題率可達 30%，故最佳終止條件值為 0.7。

關鍵詞：認知診斷的電腦化適性測驗、認知診斷、電腦化動態評量、整數四則運算

Abstract

Traditional paper and pencil tests can only know the questions of right and wrong, but students can not understand the concept of cognitive lacking, so to produce a cognitive diagnosis, result in the student not give hints or feedback immediately when the wrong answer, if the diagnosis of cognitive and effort to combine a computer adaptive dynamic assessment will be able to save time.

The research simulated the average Q matrix for each attribute concept questions 1.2, 1.8, 2.4, 3.6 and 52 Q empirical data matrix 0.7, 0.75, 0.8, 0.85, 0.9 terminate these

five conditions, results The average recognition rate and the provincial title attribute Rate this situation, thereby establishing teaching system.

The results obtained from the study: Simulate different Q matrix and the empirical data of its termination condition recognition rate for the average property assessment index as a great influence, and the empirical data in the termination condition is 0.7, the title up to 30% provincial rate, so the best termination condition is 0.7.

Keywords:

Cognitive Diagnostic Computer Adaptive Testing
Cognitive Diagnosis
Computerized Dynamic Assessment
Integer Arithmetic

1. 前言

近年來，測驗不再只看題目的對錯，更重要的是能知曉學生的能力以及他欠缺的認知概念，因此 de la Torre 在 2009 年認為認知診斷模式 (cognitive diagnosis models, CDMs) 是根據受試者是否精熟試題所需之概念屬性而形成答對機率的模式 (江鴻鈞, 2012)，而此模式中又以 DINA 模式 (Deterministic Inputs, Noisy "and" Gate Model) 最為容易 (江鴻鈞, 2012)，但為了使測驗適性化，最重要莫過於選題法。目前選題法種類眾多，最為常見的有 Kullback-Leibler (KL 法)、Shannon Entropy (SHE 法)、Posterior-weighted KL information (PWKL 法) 以及 Hybrid Kullback-Leibler (HKL 法)，其中 Kullback-Leibler 訊息的後驗加權 KL 法 (Posterior-weighted KL information, PWKL)

其演算方式又是以 Attribute Discrimination Index A (ADI) 的公式做為選題的方法，因此 ADI 與學生估計認知概念組型有著密切之關係 (Cheng, 2009; 江鴻鈞, 2012)，故本研究選擇 PWKL 做為選題之方法。

雖然認知診斷模式能瞭解學生所欠缺或精熟概念屬性，但若學生人數過多，將無法針對每一位學生的概念屬性瞭若指掌，因此結合電腦化之動態評量，可使學生藉由作答題目時直接做補救教學，降低教師課後輔導之困難，如此一來測驗方式就會變為「選題—測驗—提示—再測驗—再提示—終止提示—再選題」，此方式可使學生發現錯誤的概念，進而導正觀念，不僅縮短測驗時間且客觀。

國小數學教材則是由教育部在 2003 年訂定，且「整數四則運算」在國小學習「數與計算」中占有重要的位置 (徐偉民、林潔慧, 2010)。

2. 目的

綜合上述本研究是以認知診斷 DINA 模式為基礎建置電腦化適性動態評量教學系統，而試卷則是以國小四年級整數四則運算為例。

本研究之主要目的如下：

- 2.1 探討不同 Q 矩陣在不同終止條件下，其本研究平均屬性辨識率與省題率。
- 2.2 以國小數學整數四則運算實證資料驗證模擬實驗的結果。
- 2.3 建置以 DINA 模式為基礎之電腦化適性動態評量教學系統。

3. 文獻探討

3.1 DINA 模式

DINA 模式是假設學生具備試題所需具有之認知概念時，在理想狀態下就代表會答對該題，也就 $x_i = 1$ 是 $x_i = 1$ ；反之，缺乏任一個需要的概念時，則會答錯試題，也就是 $x_i = 0$ ，然而實際上答對機率是會受到粗心 (slipping) 與猜測 (guessing) 兩參數所影響 (卓淑瑜, 2011; 白曉珊、楊智為、郭伯臣、陳俊華, 2012; 楊智為、郭伯臣、吳慧珉、江鴻鈞, 2012)，而 DINA 選題法則是由 Cheng 在 2009 年改良 KL 法且提出基於 Kullback-Leibler 訊息的後驗加權 KL 法 (posterior-weighted KL, PWKL)，其理念則是將每個潛在認知概念的組型加入其後驗分布作為權重 (江鴻鈞, 2012)。

3.2 電腦化動態評量

動態評量即在測驗進行當中，給予學生教學上提示、線索及協助，進而獲得學生「最大可能操作水準」的訊息 (邱上真, 1996; 黃淑津、鄭麗玉, 2004)。動態評量具有六種模式，分別為「測驗—訓練—測驗」模式、習潛能評量模式或中介評量、漸進提示評量模式、連續評量模式、測量極限評量、心理計量模式，本研究是採漸進提示評量模式，當學生答錯時漸進提示會給學生本身所需要的教學提示，且提示會因學生答錯的次數而給予不同的深度提示，簡單來說漸進提示會因學生之能力而有所不同。

3.3 電腦化適性動態評量

de la Torre 在 2009 年認為認知診斷模式 (cognitive diagnosis models, CDMs) 是根據學生是否精熟試題所需之概念屬性而形成答對機率的模式，此模式需要建立關聯矩陣 (incidence matrix) 也就是 Q 矩陣則，Q 矩陣是指試題與認知屬性的關係，(incidence matrix) (Tatsuoka, 1985)，其結構是由數值 0 與 1 所組成，表示試卷中的試題所測量的特定概念，如有 j 個試題與 K 個概念，則 Q 矩陣的大小為 $J \times K$ ， q_{jk} 代表要解決第 j 個試題是否需具備概念 k，若需要則 q_{jk} 為 1，反之則為 0 (江鴻鈞, 2012)。

認知診斷模式可以用來診斷受試者是否具備測驗所需的認知概念，其認知診斷模式是利用一個潛在向量 $\alpha_i = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{ik})$ 來表示受試者是否精熟每一個認知概念，其中 $\alpha_{ik} = 1$ 表示第 i 位受試者精熟第 k 個認知概念， $\alpha_{ik} = 0$ 則表示未精熟第 k 個認知概念 (Tatsuoka, 1985; 涂金堂, 2003; 楊智為、卓淑瑜、郭伯臣、陳亭宇, 2011; 卓淑瑜, 2011)。

3.4 整數四則運算

「整數四則運算」在國小「數與計算」中占有中樞之位置且是四年級數學課程中較為基礎的單元之一 (徐偉民、林潔慧, 2010)，「整數四則運算」是根據教育部訂定五大能力指標，分別為表 3-4-1：

編碼	分段能力指標
N-2-03	能熟練整數加、減的直式計算。
N-2-05	能理解乘、除直式計算。
N-2-06	能在具體情境中，解決兩步驟問題(含除法步驟)。
N-2-07	能做整數四則混合運算，理解併式，

	並解決生活中的問題。
A-2-02	能在具體情境中，理解乘法結合律，並運用於簡化計算。

表 3-4-1 教育部訂定五大能力指標

而後由康軒文教集團其四年級團隊設計出102年四年級第二學期數學領域之教學計畫表，此教學計畫表為五大能力指標之分年細目，如表3-4-2：

編碼	分年細目
4-n-03	能在具體情境中，解決兩步驟問題，並學習併式的記法（包括連乘、連除、乘除混合）。
4-n-04	能作整數四則混合計算（兩步驟）。
4-a-01	能在具體情境中，理解乘法結合律、先乘再除與先除再乘的結果相同，也理解連除兩數相當於除此兩數之積。

表3-4-2 102年康軒版本四年級第二學期數學領域之教學計畫表

最後再由本研究者與一位專業教師共同設計出子技能與概念屬性，如表3-4-3、3-4-4：

編碼	子技能
S01	能理解橫式計算方式是由左而右計算。
S02	能理解括號內的算式必須先運算。
S03	能理解結合律的運算規則。
S04	能理解先乘除後加減的運算規則。
S05	能熟練整數四則運算的併式記法。
S06	能理解連除兩數相當於除此兩數之積。

表3-4-3本研究者與專業教師共同設計之子技能

編碼	概念屬性
A01	能理解橫式計算方式是由左而右計算。
A02	能清楚括號內的算式必須先運算。
A03	能運用乘法結合律讓計算過程簡化。
A04	熟練先乘除後加減的運算規則。
A05	能理解題意，並列成一個算式。
A06	能理解連除兩數相當於除此兩數之積。

表 3-4-4 本研究者與專業教師共同設計之概念屬性

4. 研究方法

4.1 研究流程

本研究是以認知診斷模型為基礎，探討兩種情形，其一為不同的屬性個數在不同的終止條件下其平均屬性辨識率作為評估指標，與省題率之分佈，其二為模擬資料與實徵資料在不同的終止條件下，其平均屬性辨識率與省題率之差異。透過軟體進行模擬樣本資料的產生，使用DINA模式之pwkl選題方法進行估計終止條件，探討估計終止條件的準確性與穩定性。

4.2 模擬資料

本研究使用 Matlab (8.1.0.604) 版本進行資料模擬，並估計其屬性的平均屬性辨識率。Q 矩陣先分為每題平均 1.2 個、1.8 個、2.4 個、3.6 個等四種不同的屬性個數，再由 Matlab 隨機產生 50 筆且人數為 1000 人、題庫為 300 題、slip 與 guess 上下限皆為 0.05~0.25 之間的作答反應，最後再分別設 0.7、0.75、0.8、0.85、0.9 這五種終止條件，是為了瞭解每題平均不同個數之屬性與不同終止條件下，其平均屬性辨識率與省題率之分佈，整理如表 4-2-1

Q 矩陣每題平均屬性	1.2	1.8	2.4	3.6
題庫	300	300	300	300
人數	1000	1000	1000	1000
guess	0.05~0.25	0.05~0.25	0.05~0.25	0.05~0.25
	5	5	5	5
slip	0.015~0.475	0.015~0.475	0.015~0.475	0.015~0.475
	5	5	5	5
測驗題數	18	18	18	18
終止	0.7	0.7	0.7	0.7
	0.75	0.75	0.75	0.75
	0.8	0.8	0.8	0.8

條件值	0.85	0.85	0.85	0.85
	0.9	0.9	0.9	0.9

表 4-2-1 不同平均屬性在不同的終止條件的試題參數

4.3 實徵資料

4.3.1 研究對象

本研究以台中市南區信義國民小學五年級2班，共52名學童為研究的實徵對象，採紙筆測驗方式進行施測。

4.3.2 研究題目

本研究選擇單元是以教育部頒定國小數學教材「數與計算」中，四年級第二學期「整數四則運算」為例，而試卷題目則是由本研究與一位專業教師共同命題。

4.4 評估準則

4.4.1 辨識率

本研究使用辨識率作為模擬資料之評估標準。辨識率意指模擬產生之受試者其認知組型與模式估計之概念組型，其兩者之間的相符程度（江鴻鈞，2012）。

辨識率計算公式 (1) 如下：

$$\text{辨識率} = \frac{\text{模式估計與模擬資料真值相同的樣本數}}{\text{全部樣本數}} \quad (1)$$

4.4.2 分類的一致性 (consistency of classification index, CCI)

本研究使用分類的一致性 (consistency of classification index, CCI) 作為評估標準。一致性意指受試者的認知屬性狀態在認知診斷模式之估計是否與專家判定的結果一致，一致性愈高，表示認知診斷模式之估計成效越好。

一致性計算公式 (2) 如下：

$$CCI = \frac{\text{模式估計與專家判斷相同的樣本數}}{\text{全部樣本數}} \quad (2)$$

4.4.3 節省題數率

本研究使用節省題數率作為模擬資料之評估標準。省題率意指受試者的平均作答題數佔總題數的百分之多少。

省題率計算公式 (3) 如下：

$$\text{省題率} = (\text{總題數} - \text{平均施測題數}) \div \text{總題數} \quad (3)$$

5. 研究結果

5.1 模擬之研究結果

研究結果發現，在 DINA 模式下，不同 Q 矩陣在不同終止條件其平均屬性辨識率與省題率，如表 5-1：

Q 矩陣每題平均屬性	終止條件值	guess=0.05~0.25 slip=0.05~0.25	
		平均屬性辨識率	省題率
1.2	0.7	0.950867	0.5486944
	0.75	0.960167	0.4962111
	0.8	0.968793	0.4384556
	0.85	0.977567	0.3750389
	0.9	0.985707	0.3069333
1.8	0.7	0.951797	0.5234000
	0.75	0.961200	0.4771278
	0.8	0.969207	0.4258333
	0.85	0.977910	0.3689333
	0.9	0.985450	0.3074278
2.4	0.7	0.951150	0.5173778
	0.75	0.960380	0.4727056
	0.8	0.969303	0.4234611
	0.85	0.977480	0.3676722
	0.9	0.984933	0.3080389
3.6	0.7	0.949503	0.4672278
	0.75	0.958597	0.4215944
	0.8	0.967247	0.3719889
	0.85	0.975037	0.3167833
	0.9	0.981683	0.2573000

表 5-1 不同 Q 矩陣在不同終止條件的平均屬性辨識率與省題率

5.2 實徵資料在模擬中其研究結果

在 DINA 模式下，國小數學整數四則運算實證資料在模擬中其分類的一致性省題率，如表 5-2：

實徵資料	終止條件值	CCI	省題率
5 年級 2 個班共 52 人	0.7	0.9744	0.3023500
	0.75	0.9808	0.0416667
	0.8	0.9808	0.0096167
	0.85	0.9808	0.0096167
	0.9	0.9776	0.0096167

表 5-2 實證資料在模擬中分類的一致性與省題率



圖 5-3-4 試卷分配介面

6. 結論

6.1. 由表 5-1 可看出兩個部分，其一為不同的 Q 矩陣其終止條件皆界在 0.7~0.75 為最佳終止條件；其二為當 Q 矩陣每題平均屬性越多時，省題率越低，其中以 Q 矩陣每題平均有 1.2 個屬性其省題率最高，因此概念屬性在每題中能夠被清楚界定，其省題率會最佳。

6.2. 在實徵資料驗證上，從 5-2 表可發現終止條件在 0.7 時，其分類的一致性為 0.9744 且省題率可以節省 30% 的題數，因此終止條件之設定對於屬性正確分類率影響很大，且將終止條件設為 0.7，則是最佳效益。

6.3. 由 6.2 之數據顯示最佳終止條件為 0.7，因此將最佳終止條件寫入程式，並建置電腦化適性動態評量教學系統，進而使教學系統達到分類的一致性與省題率。

綜合上述，本研究依據探討之結果，將終止條件值寫入程式內並建置教學系統，未來可針對此教學系統與無終止條件值進行比較。

謝誌

本研究感謝國科會之經費補助，計畫編號 NSC 102-2511-S-142-008-MY3。

參考文獻

- [1] 黃淑津、鄭麗玉，電腦化動態評量對國小五年級學生閱讀理解效能之研究，*國立嘉義大學國民教育研究所國民教育研究學報*第 12 期，pp.167-201，2004。
- [2] 林潔慧、徐偉民，利用教學模組進行國小四年級四則運算兩步驟文字題補救教學之行動研究，*屏東教育大學學報-教育類*第三十四期，pp. 211-242，2010。
- [3] 卓淑瑜，*不同認知診斷適性測驗演算法結合*

知識結構之成效比較，國立臺中教育大學教育測驗統計研究所碩士論文，2011。

- [4] 卓淑瑜、郭伯臣、陳亭宇、楊智為，DINA 與 G-DINA 模式參數不變性探討，*國立臺中教育大學教育測驗統計研究所測驗統計年刊*第十九輯，2011。

- [5] 江鴻鈞，*基於 CDI 結合概念加權之認知診斷適性測驗選題法*，國立臺中教育大學教育測驗統計研究所博士論文，2012。

- [6] 白曉珊、郭伯臣、陳俊華、楊智為，認知診斷測驗選題法之模擬研究，*全球華人計算機教育應用大會*，pp.324-329，2012。

- [7] 楊智為、郭伯臣、吳慧珉、江鴻鈞 (2012)。基於 CDI 結合概念加權之認知診斷適性測驗選題法。*TWELF 2012 第八屆台灣數位學習發展研討會*。2012 年 10 月 26-27 日，國立成功大學。