

## 學習曲線分析於電腦化動態評量之應用

郭伯臣<sup>1</sup> 林佳樺<sup>2</sup> 葉珮玉<sup>3</sup> 楊智為<sup>4</sup>

### 摘要

學習曲線主要是用來描述學生學習的次數與答錯率之間的關係，透過學習曲線分析可以幫助教師了解學生的學習狀況，從而更有效的調整教學教材、進行教學，以及提供個別化的指導，因此學習曲線分析在教育領域中是相當重要的。目前，已有多種應用於學習曲線分析的統計模式被提出，要如何從多種模式中選出最貼近學生實際學習情況的模式，以提高分析效率與精準度，為本研究之目的。本研究使用較常見的統計模式（AFM、iAFM、BKT、iBKT），透過模擬方式產生不同的學習曲線，並分析實徵資料，進行模式比較，選擇出最適合的統計模式，並提出使用建議。

研究結果顯示，AFM 及 iAFM 的估計精準度比 BKT 及 iBKT 高，AFM 及 iAFM 的估計誤差值相近，而 iAFM 可以估計學生個人的學習率，更能讓教師針對不同學生的學習狀況加以輔導，提高其學習成效。此外，不同的模式表現不會受到起點行為所影響；但在理想的學習曲線中，模式表現會受到學習率的影響。本研究亦透過電腦化動態評量之實徵資料進行驗證與分析，並做為後續研究者或教師在學習資料分析上的一個範例。

**關鍵詞：**學習曲線、AFM、iAFM、BKT、iBKT

- 
1. 郭伯臣，國立臺中教育大學教育資訊與測驗統計研究所卓越講座教授
  2. 林佳樺，國立臺中教育大學教育資訊與測驗統計研究所博士後研究員
  3. 葉珮玉，國立臺中教育大學教育資訊與測驗統計研究所碩士
  4. 楊智為，國立臺中教育大學教育資訊與測驗統計研究所助理教授
- 收件日期：2023.08.22；完成修改：2023.12.11；正式接受：2024.01.07  
通訊作者：林佳樺；Email：linch2812@gmail.com  
地址：403514 臺中市西區民生路 140 號  
國立臺中教育大學教育資訊與測驗統計研究所

## Application of Learning Curve Analysis to Computerized Dynamic Assessments

Bor-Chen Kuo<sup>1</sup> Chia-Hua Lin<sup>2</sup>  
Pei-Yu Ye<sup>3</sup> Chih-Wei Yang<sup>4</sup>

### Abstract

Learning curve is applied to describe the relationship of the number of opportunity and the error rate of student responses. Learning curve analysis can assist educators in understanding students' learning situations, thus enabling more effective adjustments to teaching materials, conducting instruction, and providing personalized guidance. The purpose of this research aimed how to select the best model matching the actual learning situation of students' and improving the efficiency and accuracy of the analysis. There were four statistical models, AFM, iAFM, BKT, and iBKT, employed to generate various learning curves through simulation, in addition, our study analyzed actual data to decide the optimal statistical model among them.

The research findings indicate that AFM and iAFM exhibit higher estimation accuracy compared to BKT and iBKT. The estimation errors of AFM and iAFM are similar; however, iAFM has the capability to estimate individual student learning rates, enabling teachers to provide tailored guidance based on varying student learning conditions and enhancing overall learning outcomes. Furthermore, the performance of different models is not influenced by students' initial ability; however, in an ideal learning curve, model performance is influenced by students' learning rates. Furthermore, this study validates and analyzes through empirical data from computerized dynamic assessment, serving as an exemplar for subsequent researchers or educators in the analysis of learning data.

**Keywords:** learning curve, AFM, iAFM, BKT, iBKT

1. Bor-Chen Kuo, University Chair Professor, Graduate Institute of Educational Information and Measurement, National Taichung University of Education
  2. Chia-Hua Lin, Postdoctoral Project Researcher, Graduate Institute of Educational Information and Measurement, National Taichung University of Education
  3. Pei-Yu Ye, Master, Graduate Institute of Educational Information and Measurement, National Taichung University of Education
  4. Chih-Wei Yang, Assistant Professor, Graduate Institute of Educational Information and Measurement, National Taichung University of Education
- Received: 2023.08.22; Revised: 2023.12.11; Accepted: 2024.01.07  
Corresponding Author: Chia-Hua Lin; Email: lynch2812@gmail.com  
Address: No. 140, Minsheng Rd., West Dist., Taichung City 403514, Taiwan  
Graduate Institute of Educational Information and Measurement, National Taichung University of Education

## 壹、前言

學習分析（learning analytics）是透過學生的學習歷程來進行評估、解釋和預測，目的是在了解其學習狀態。而在教學上，教師若能運用學習分析的潛在價值，則能清楚了解每位學生的學習狀況，對其進行更有效的個別輔導，進而提升學習成效（Wang, 2011）。

Newell 與 Rosenbloom（1981）提出學習曲線模式來改善認知模式，透過錯誤率（error rate）和測驗次數來描述學習狀態。隨著概念的測驗次數增加，錯誤率會呈現指數遞減，其繪出之曲線則稱為學習曲線（learning curve）。教師可以透過學習曲線觀察到是否存在複雜概念，並經由概念的拆解來幫助學生學習，故學習曲線可以協助教師了解學生在概念上的學習狀況，並調整概念內容、概念學習順序，以及教學策略（Goldin et al., 2016; Koedinger et al., 2016）。

透過學習因素分析（learning factors analysis [LFA]; Cen et al., 2006）方法中的統計模式，能夠建構出最適合學習資料的學習曲線。目前，已有許多統計模式被提出，例如：Cen（2009）提出的 AFM 模式（additive factor model [AFM]），這也是最廣泛使用的一種學生學習模式，可估計其先備知識、概念難度，以及學習這些概念的速度。此外，還有具有個人學習率的 iAFM 模式（individualized-slope additive factor model [iAFM]; Liu & Koedinger, 2015）；考慮學生的猜測及粗心率的 BKT 模式（Bayesian knowledge tracing models [BKT]; Corbett & Anderson, 1994），以及考慮學生個人參數的 iBKT 模式（individualized Bayesian knowledge tracing models [iBKT]; Yudelson et al., 2013）等多種統計模式。

若能從多種模式中選擇出最能符合實際情況的模式，則可提高分析的效率及精準度，並進行有效的評估。但在過去的文獻中，大多係利用實徵資料進行分析，而沒有將四種模式進行比較，因此本研究修改並延伸葉珮玉（2020）之研究，比較學習曲線分析方法中，AFM、iAFM、BKT 及 iBKT 四種統計模式的估計準確度，以便分析學生的學習歷程、預測學生的表現，更進一步探討學生的起點行為、學生學習率是否會影響統計模式的預測，最後選擇出與學生學習情況最適配的模式。此外，本研究亦蒐集學生於線上學習平臺進行電腦化動態評量之學習資料，並透過實徵資料進行驗證與分析。

## 貳、文獻探討

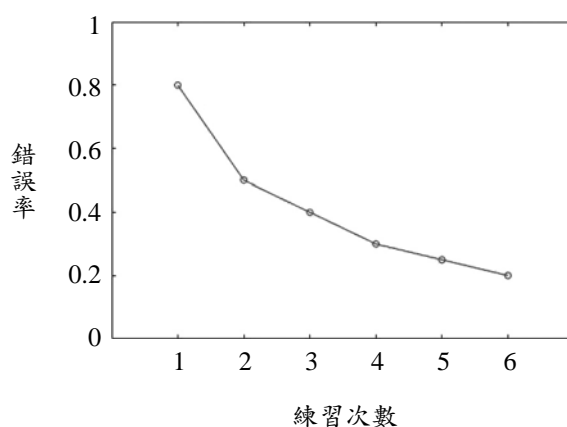
### 一、學習曲線

認知模式（cognitive model; Cen et al., 2006）主要是用來描述學生如何運用哪些概念或技能來解決問題，通常是由學科與評量專家來定義，然而有許多研究顯示，專家定義的認知模式經常錯誤的反應學生之學習狀態（Koedinger & Nathan, 2004; Nathan & Koedinger, 2000a, 2000b; Nathan & Petrosino, 2003; Nathan et al., 2002）。因此，Newell 與 Rosenbloom（1981）提出學習曲線模式來改善認知模式，透過函數  $Y = aX^b$  來描述學習曲線，其中  $Y$  為錯誤率（error rate）， $X$  為概念的練習次數， $a$  為概念的難易度， $b$  為學習率（學習概念的難易程度）。學習曲線顯示學生在學習的過程中，行為變化累積之效果，依特性可將學習曲線圖分為四種類型，說明如下。

#### （一）理想學習曲線

錯誤率經一定的練習次數，會隨著練習次數增加而下降（Koedinger et al., 2013），是學習曲線最理想的狀態（如圖 1 所示）。

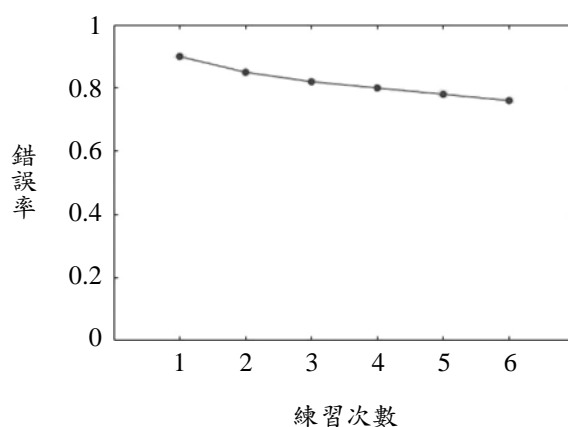
圖 1 理想學習曲線圖



## （二）高平學習曲線

高平（still high）學習曲線表示經多次練習後，總體錯誤率仍是偏高的情形（DataShop, n.d.），其原因有兩種：第一是學生沒有足夠的練習而無法掌握該技能；第二是試題可能含有兩種以上的混合技能，此較為複雜（Corbett & Anderson, 1994）（如圖 2 所示）。

圖 2 高平學習曲線圖



## （三）低平學習曲線

低平（low and flat）學習曲線的總體錯誤率一開始是偏低（DataShop, n.d.），經多次練習後，錯誤率依舊維持低平，表示學生已掌握該技能，或是該試題過於簡單，學生係透過猜測及簡單的答題策略來獲得正確答案（Goldin et al., 2016）（如圖 3 所示）。

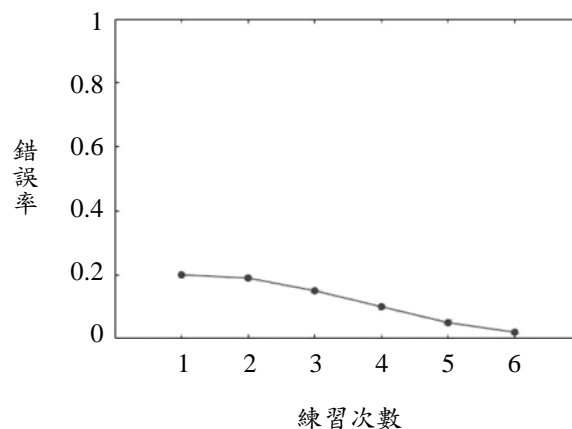
## （四）學習無效學習曲線

學習無效（no learning）學習曲線與低平學習曲線的差異，在於其總體錯誤率下降得更為緩慢（學生之學習率趨近於 0）（DataShop, n.d.），表示此技能可能較具有挑戰性（Goldin et al., 2016）。

Corbett 與 Anderson（1994）觀察到，在一些複雜概念中的指數關係可能

測驗學刊，第 71 輯第 3 期

圖 3 低平學習曲線圖



並不明顯，此學習曲線上會有突然出現高出的尖點（blip），如圖 4 所示。如果複雜概念（S）可以分解成單一概念（S1、S2、S3），那麼指數關係就會成立，且每個單一概念都會呈現出更加平滑的學習曲線，如圖 5 所示。概念 S1 一開始學習時的錯誤率較低，且經過幾次練習後，大部分學生都可掌握此概念，故教師可依拆解後的單一概念來調整教學內容與順序，來幫助學生學習。葉珮玉等人（2019）針對電腦化動態評量之學習資料，將概念拆解成以能力指標、子技能、試題為單位的三種不同之 KC 模式（knowledge component model），研究結果顯示以試題為單位的模式適配度較佳。

圖 4 複雜概念的學習曲線

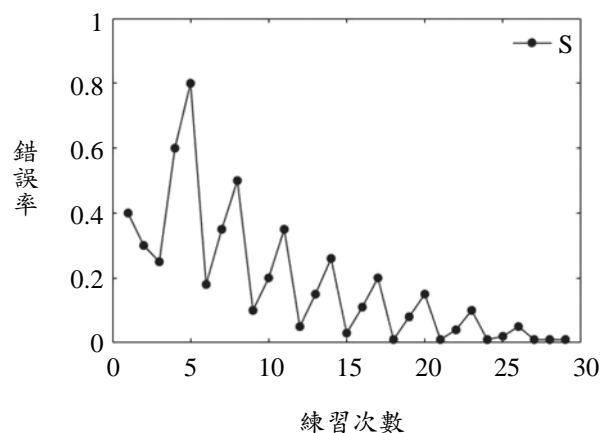
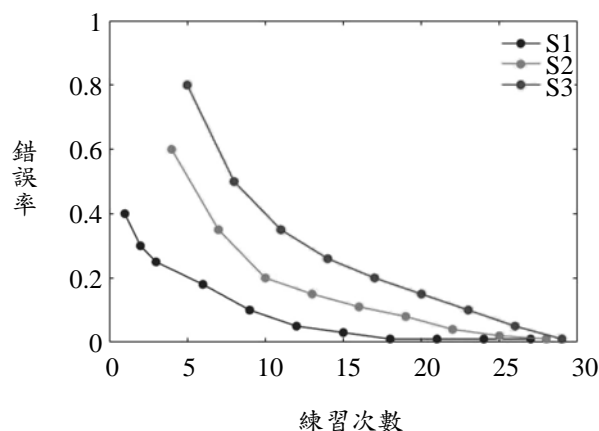


圖 5 切割的子概念之學習曲線



## 二、學習曲線分析之統計模式

常用來建構學習曲線的統計模式有 AFM、iAFM、BKT 及 iBKT 等，近年來也有一些基於 AFM 或 BKT 所衍生的模式，例如：運用於有操作或解題歷程的部分計分模式 PC-AFM (Rachatasumrit & Koedinger, 2021)，研究結果顯示 PC-AFM 優於 AFM。另外，也有結合 IRT 及 BKT 兩種模式的 IRT-BKT (Wang et al., 2017)，同時估計 IRT 的學生能力值與試題參數，以及 BKT 的學習、猜測與遺忘參數，研究發現 IRT-BKT 的參數估計精準度優於 IRT 及 BKT。

考量現今電腦化評量多為二元計分，且本研究所蒐集之實徵資料亦為二元計分，另考量本研究著重於探討個人化學習參數的影響，故以下僅針對 AFM 及 BKT 之原型模式，及其擴展之 iAFM 及 iBKT 進行說明與比較。

### (一) AFM 模式

AFM 模式 (AFM; Cen, 2009) 是一種認知診斷模型，可估計學生在學習前的初始能力、概念難易度，以及概念學習率，其數學公式定義如下：

$$P(X_{ij} = 1 | \theta_i, \beta, \gamma) = \frac{\exp(\theta_i + \sum_{k=1}^K Q_{jk}(\beta_k + \gamma_k T_{ik}))}{1 + \exp(\theta_i + \sum_{k=1}^K Q_{jk}(\beta_k + \gamma_k T_{ik}))} \quad (1)$$

其中， $\theta_i$  為第  $i$  位學生的初始能力， $\beta_k$  為第  $k$  個概念的難度係數， $\gamma_k$  為第  $k$  個



測驗學刊，第 71 輯第 3 期

概念的學習率， $T_{ik}$  為第  $i$  位學生在第  $k$  個概念的練習次數。

$X_{ij}$  為第  $i$  位學生在第  $j$  題試題的作答反應，此作答反應為二元計分型態， $X_{ij}=0$  表示作答錯誤、 $X_{ij}=1$  表示作答正確； $Q_{ik}=1$  表示第  $j$  題用到第  $k$  個技能， $Q_{ik}=0$  表示第  $j$  題沒有用到第  $k$  個技能，其中  $\theta_i$  值愈小表示學生的初始能力愈低； $\beta_k$  值愈小表示概念  $k$  難度愈高； $\gamma_k$  愈小表示概念  $k$  愈不容易學會； $T_{ik}$  值愈小表示學生在概念  $k$  的練習次數愈少。因此， $\theta_i$ 、 $\beta_k$ 、 $\gamma_k$  或  $T_{ik}$  值愈小，表示學生在試題  $j$  上的答對機率愈低。

## (二) iAFM 模式

iAFM 模式 (iAFM; Liu & Koedinger, 2015) 係以 AFM 為基礎，增加學生個人的學習率參數，以提升模式參數估計值的信度與效度。其數學公式定義如下：

$$P(X_{ij}=1 | \theta_i, \beta, \gamma, \delta) = \frac{\exp(\theta_i + \sum_{k=1}^K Q_{jk}(\beta_k + \gamma_k T_{ik} + \delta_i T_{ik}))}{1 + \exp(\theta_i + \sum_{k=1}^K Q_{jk}(\beta_k + \gamma_k T_{ik} + \delta_i T_{ik}))} \quad (2)$$

其中， $\theta_i$  為第  $i$  位學生的初始能力， $\beta_k$  為第  $k$  個概念的難度係數， $\gamma_k$  為第  $k$  個概念的學習率， $T_{ik}$  為第  $i$  位學生在第  $k$  個概念的練習次數， $\delta_i$  為第  $i$  位學生的個人學習率。 $\delta_i$  值愈小，表示學生的個人學習效果愈不佳。

## (三) BKT 模式

BKT 模式 (BKT; Corbett & Anderson, 1994) 是一種廣泛應用於教育資料探勘與學習分析領域的機率模型 (Bulut et al., 2023; David et al., 2016; Sao Pedro et al., 2013)，也是一種隱藏式馬可夫模式 (Hidden Markov Model [HMM])，藉由此模式能夠了解學生是否已掌握特定概念，而其每次的學習表現都會因為學習率而有所更動。

BKT 模式主要有四個參數，分別為： $P(L_{j0})$  表示學生在尚未學習前，掌握第  $j$  個概念之機率； $P(G_j)$  表示學生在沒有掌握第  $j$  個概念時，猜對之機率； $P(S_j)$  表示學生在掌握第  $j$  個概念時，粗心之機率； $P(T_j)$  表示學生學習第  $j$  個概念之學習率。其數學公式如下：

$$P(L_{jt}) = P(L_{jt-1}) + P(T_j)(1 - P(L_{jt-1})) \quad (3)$$



假設學生在第  $j$  個概念的練習次數為  $t$  次， $P(L_{jt})$  為學生在第  $j$  個概念練習  $t$  次後所掌握該概念之機率。根據公式 3，可以從前一次的答對機率和學生的學習率來預測出下一次掌握概念之機率。

公式 4 的  $P(C_{jt})$  為學生在第  $j$  個概念練習  $t$  次後，實際掌握第  $j$  個概念之機率，在真實的學習環境中，學生有可能會有粗心答錯與猜測答對的情況，因此學生實際掌握概念的機率如公式 4。

$$P(C_{jt}) = P(G)(1 - P(L_j)) + (1 - P(S_j))P(L_j) \quad (4)$$

公式 5 及公式 6 的  $O_t = \{o_1, \dots, o_t\}$  為學生在第  $t$  次的作答反應，已知學生的作答反應，計算出學生在第  $t$  次練習後所掌握第  $j$  個概念之機率。若學生作答正確 ( $O_{t=correct}$ )，則以公式 5 計算出  $P(L_{jt-1})$ ；若學生作答錯誤 ( $O_{t=incorrect}$ )，則以公式 6 計算出  $P(L_{jt-1})$  (Baker et al., 2008a)。

$$P(L_{jt-1} | O_{t=correct}) = \frac{P(L_{jt-1})(1 - P(S))}{P(L_{jt-1})(1 - P(S)) + (1 - P(L_{jt-1}))P(G)} \quad (5)$$

$$P(L_{jt-1} | O_{t=incorrect}) = \frac{P(L_{jt-1})rmP(S)}{P(L_{jt-1})P(S) + (1 - P(L_{jt-1}))(1 - P(G))} \quad (6)$$

#### (四) iBKT 模式

因有許多研究顯示，考慮學生的個別因素可增加模式的準確度 (Baker et al., 2008b; Cen et al., 2008; Pardos & Heffernan, 2010)，故 Yudelson 等人 (2013) 提出了 iBKT 模式，與 BKT 最大差異在於 iBKT 能估計學生個人的概念掌握機率  $P(L)$ 、個人的猜測率  $P(G)$ 、個人的粗心率  $P(S)$ ，以及個人的學習率  $P(T)$ 。公式如下：

$$P(L_{ijt}) = P(L_{ijt-1}) + P(T_{ij})(1 - P(L_{ijt-1})) \quad (7)$$

$$P(C_{ijt}) = P(G_{ij})(1 - P(L_{ijt})) + (1 - P(S_{ij}))P(L_{ijt}) \quad (8)$$

$$P(L_{ijt-1} | O_{ijt=correct}) = \frac{P(L_{ijt-1})(1 - P(S_{ij}))}{P(L_{ijt-1})(1 - P(S_{ij})) + (1 - P(L_{ijt-1}))P(G_{ij})} \quad (9)$$

$$P(L_{ijt-1} | O_{ijt=incorrect}) = \frac{P(L_{ijt-1})P(S_{ij})}{P(L_{ijt-1})P(S_{ij}) + (1 - P(L_{ijt-1}))(1 - P(G_{ij}))} \quad (10)$$

其中， $P(L_{ij0})$  表示第  $i$  位學生在尚未學習前，掌握第  $j$  個概念之機率； $P(G_{ij})$  表示第  $i$  位學生在沒有掌握第  $j$  個概念時，猜對之機率； $P(S_{ij})$  表示第  $i$  位學生在掌握第  $j$  個概念時，粗心之機率； $P(T_{ij})$  表示第  $i$  位學生學習第  $j$  個概念之學習率； $P(L_{ijt})$  為第  $i$  位學生在練習技能  $t$  次後，能掌握第  $j$  個概念機率； $P(C_{ijt})$  為第  $i$  位學生練習技能  $t$  次後，實際上答對第  $j$  個概念之機率； $O_{ijt} = \{o_{ijt1}, \dots, o_{ijtT}\}$  為第  $i$  位學生在第  $j$  個概念的第  $t$  次作答反應，已知第  $i$  位學生的作答反應，計算出學生在第  $t$  次練習後所掌握該概念之機率，若第  $i$  位學生作答正確，則以公式 9 計算出  $P(L_{ijt-1})$ ；若第  $i$  位學生作答錯誤，則以公式 10 計算出  $P(L_{ijt-1})$ 。計算出  $P(L_{ijt-1})$  後，就可以從公式 8 推算出第  $i$  位學生在第  $j$  個概念練習  $t$  次後，實際上答對第  $j$  個概念之機率。

### 三、四個模式的差異與關聯

AFM 模式採用邏輯迴歸 (logistic regression) 來估計參數和預測學生的表現，並假設學習是漸進的，隨著練習機會增加，學生正確答對的機率也會增加。AFM 常用於分解認知過程、預測行為表現與改善課程設計等。Durand 等人 (2017) 透過 AFM 分析實徵資料，探討觀察值增加及減少、概念個數、學生能力是否會影響 AFM 的估計效果，並根據過去的測驗成績、其相關的評估圖來預測學生未來的學習表現。Rivers 等人 (2016) 則利用 AFM 來分析學生在學習程式碼的學習歷程，研究結果顯示：可透過 AFM 分析來改善教學及調整未來課程設計。

BKT 模式採用 HMM 來估計學生在特定主題或技能的知識狀態，透過其先前的答題紀錄來預測下一題的答題狀況，並假設在每項技能上，學生只有知道或不知道的兩種反應，透過每次的練習機會，學生都有一定的機率習得該技能。BKT 模式常用於教育科技平臺、學習成效評估與個人化教學等。Qiu 等人 (2011) 透過 BKT 模式探討時間的間隔是否會對學生的學習有所影響，研究結果顯示：遺忘參數與粗心參數皆會受到時間所影響。

另外，也有幾篇文獻主要在探討模式比較：第一篇透過實徵資料的分析，比較 AFM 及 iAFM 與 BKT 及 iBKT，估計結果顯示，有考慮學生個人參數的 iAFM 及 iBKT 模式比原先的 AFM 及 BKT 模式好（即 iAFM 優於 AFM；iBKT 優於 BKT）(Liu & Koedinger, 2017)，文中雖有列出四種模式的模式適配度，但作者並未進行 iAFM 及 iBKT 的比較與說明，並表示兩種模式的選用應取決於研究問題，此外該研究僅針對實徵資料進行分析，分析結果可能

會有所侷限；第二篇是使用 AFM 及 BKT 模式比較不同參數預測相同的學習曲線，結果顯示 AFM 的估計成效較佳（Klingler et al., 2015）；第三篇是使用 AFM 及 BKT 模式來分析學生於數位學習平臺的學習資料，結果與前篇相同，AFM 的模式適配度較 BKT 佳（林佳樺等人，2018）；第四篇則是使用 AFM 及 iAFM 模式分析電腦化動態評量，結果顯示 iAFM 的模式適配度較 AFM 佳（Yan et al., 2019）；第五篇則是在 27 個不同型態的資料集（包含 12 種不同的課程，對象從小學到大學）中比較 AFM 及 iAFM 的模式適配度，結果顯示 iAFM 與資料集的適配次數多於 AFM（Koedinger et al., 2023）。

從上述文獻可知，沒有研究同時探討 AFM、iAFM、BKT、iBKT 四個模式的估計成效，且沒有透過模擬研究探討不同的起點行為、人數與學習率等因素的影響，因此本研究修改並延伸葉珮玉（2020）之研究，利用不同的模擬資料，比較四種模式的估計成效，並探討哪些因素會影響模式估計。由於葉珮玉的研究顯示，學生人數多寡並不影響模式表現，故本篇研究並未探討不同人數的影響。

綜整四種模式的差異與關聯，表 1 針對四個模式所估計的參數進行比較，由表中可知，iAFM 較 AFM 多考慮學生個人的學習率（ $\delta_i$ ）；iBKT 不同於 BKT，在每個參數上皆以個別學生為基礎。

表 1 各模式估計參數比較表

| 模式   | 估計參數  |
|------|---|
| AFM  | $\theta_i, \beta_k, \gamma_k, T_{ik}$           |
| iAFM | $\theta_i, \beta_k, \gamma_k, T_{ik}, \delta_i$ |
| BKT  | $L_{j0}, G_j, S_j, T_j$                         |
| iBKT | $L_{ij0}, G_{ij}, S_{ij}, T_{ij}$               |

## 參、研究方法

### 一、分析工具

本研究使用的分析工具主要有 DataShop、LearnSphere，以及改編 Cloudspurs 在 GitHub 共享的程式碼做為分析。

在 AFM、iAFM 及 BKT 分析方面，使用美國自然科學基金會支持建立的

學習、科學資料庫——DataShop 與 LearnSphere。DataShop 是一個大型教育資料庫，透過網頁的形式，提供公共和私人的教育資料共享、儲存和分析的平臺，其用途很廣泛，可以提供實驗分析過程中的數據，透過分析的資料，來提升學生學習成效、預測學生表現、驗證後設認知模式等。而 LearnSphere（2018）則是由 CMU、MIT 與 Memphis 大學的研究人員開發，提供一系列教育資料分析的平臺，其優點包含有多種的分析模式可供使用者選擇，以及基於組件的流程模式，在分析、操作或可視化數據上，提供便利性高、操作簡單的使用介面。

因為 LearnSphere 尚未提供 iBKT 的分析模式，因此本研究採用 Cloudspurs（2019）在 GitHub 平臺中共享的程式碼進行修改，原先的程式碼主要是採用 BKT 模式，因 BKT 是一種 HMM，在此程式中參數估計的方法上則採用 Baum-Welch 演算法，該演算法是用於尋找 HMM 未知參數的一種 EM 算法，利用前向—後向演算法來計算 E-Step 的統計訊息。在本研究中，將原先 BKT 估計的程式碼加入學生的個人參數，進一步發展成 iBKT 的模式，以便比較四種模式的差異性。

## 二、模擬實驗設計

本研究採用模擬研究方式來探討 AFM、iAFM、BKT 及 iBKT 四種模式的估計成效，並進行比較。模擬實驗設計之學生人數為 1,000 人，技能個數為 3 個，測驗次數為 6 次，模擬出十二種不同的學習曲線圖（如圖 6～圖 9 所示），每種學習曲線皆模擬 10 次，並計算其平均誤差值，以判斷模式估計精準度。

為了探討四種模式在不同資料型態上的估計成效，本研究根據實務上可能出現的四種學習曲線型態（理想型、高平、低平、學習無效）設定學生的學習率（有效、無效）及學生不同的起始能力（高、中、低錯誤率），並模擬出 12 種不同的學習曲線圖，依據不同學習曲線所設計之錯誤率（如表 2 所示），模擬出學生的作答反應。圖 6 中的類型 A、B、C 三種皆為理想的學習曲線圖：類型 A 的學習率較高（斜率較大），因此錯誤率下降的幅度也較大；類型 B 的學習率較類型 A 低，因此在第二次測驗的錯誤率明顯高於類型 A；類型 C 的學習率較為平緩，學生需要經過多次的測驗才能夠讓錯誤率降到最低。

表 2 學習曲線模擬參數（錯誤率）設定表

| 類型 | 第一次測驗<br>(起始能力) | 第二次測驗 | 第三次測驗 | 第四次測驗 | 第五次測驗 | 第六次測驗 |
|----|-----------------|-------|-------|-------|-------|-------|
| A  | 0.95            | 0.1   | 0.025 | 0     | 0     | 0     |
| B  | 0.95            | 0.4   | 0.1   | 0.05  | 0     | 0     |
| C  | 0.95            | 0.7   | 0.45  | 0.2   | 0     | 0     |
| AL | 0.75            | 0.1   | 0.025 | 0     | 0     | 0     |
| BL | 0.75            | 0.3   | 0.075 | 0.025 | 0     | 0     |
| CL | 0.75            | 0.5   | 0.25  | 0     | 0     | 0     |
| DH | 0.9             | 0.85  | 0.825 | 0.8   | 0.75  | 0.7   |
| D  | 0.5             | 0.45  | 0.425 | 0.4   | 0.35  | 0.3   |
| DL | 0.2             | 0.15  | 0.125 | 0.1   | 0.05  | 0     |
| EH | 0.9             | 0.89  | 0.88  | 0.87  | 0.86  | 0.85  |
| E  | 0.5             | 0.49  | 0.48  | 0.47  | 0.46  | 0.45  |
| EL | 0.2             | 0.19  | 0.18  | 0.17  | 0.16  | 0.15  |

圖 6 類型 A、B、C 學習曲線圖

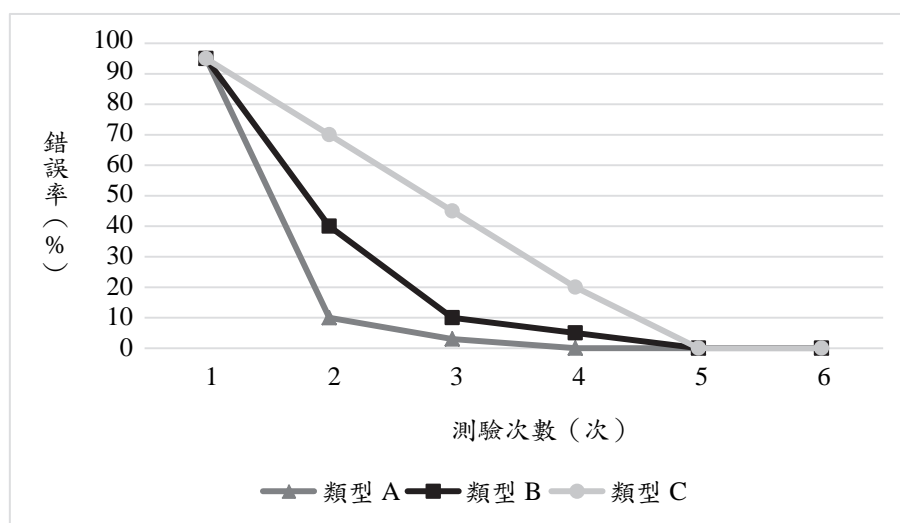


圖 7 中的類型 AL、BL 及 CL 三種亦為理想學習曲線，但類型 AL、BL 及 CL 的學生起始能力較類型 A、B 及 C 三者高，模擬此種學習曲線是為了研究學生不同的起始能力是否會造成模式上估計之差異。

測驗學刊，第 71 輯第 3 期

圖 7 類型 AL、BL、CL 學習曲線圖

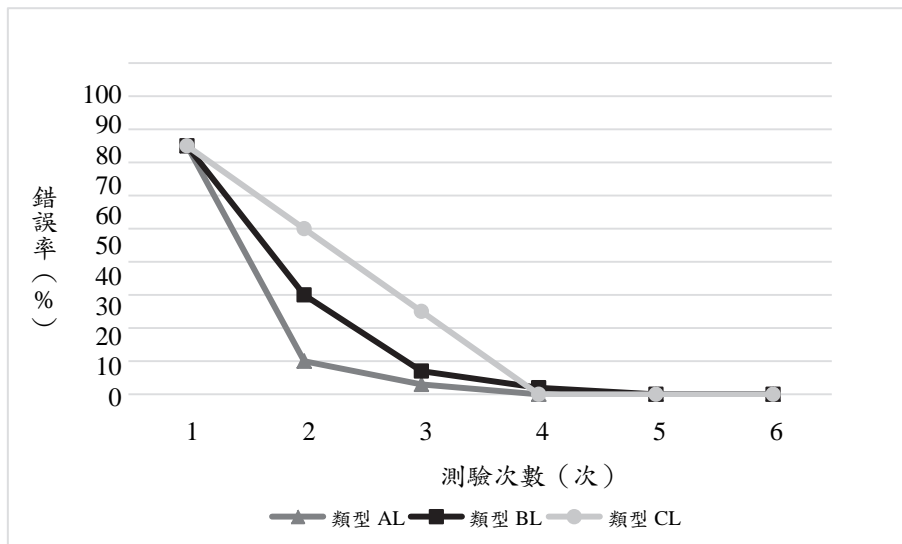


圖 8 中的類型 DH、D 及 DL 三種分別為高平、中平、低平的學習曲線圖，屬於學習緩慢的學習曲線，總體錯誤率下降趨勢緩慢。模擬此種學習曲線同樣是為了研究學生的起始點不同是否會造成模式上估計之差異，故依學生的起始能力高低順序分為類型 DH、D 及 DL。

圖 8 類型 DH、D、DL 學習曲線圖

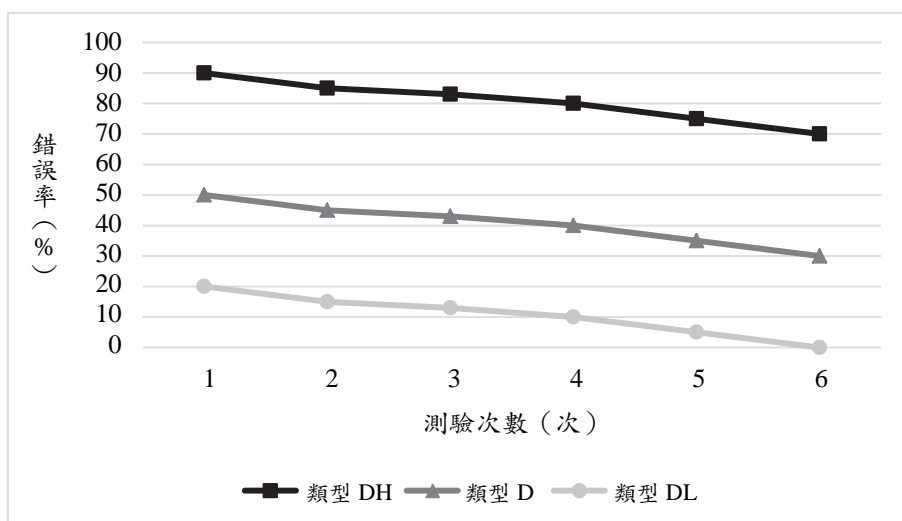
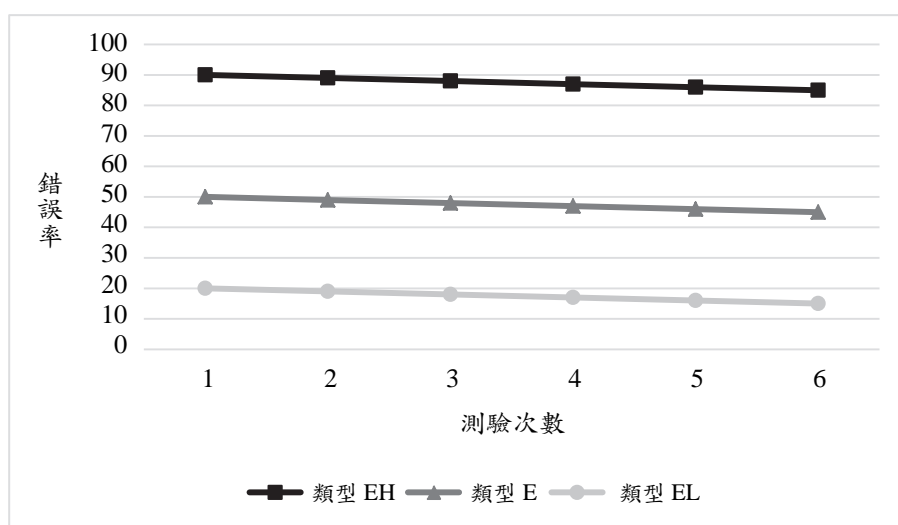


圖 9 中的類型 EH、E 及 EL 三種同樣分別為高平、中平、低平的學習曲線圖，但在總體錯誤率而言，下降的趨勢趨近於 0，表示為學習無效的學習曲線，其目的在於研究學生的起始能力差異及學習率下降幅度對於模式估計之影響。本研究模擬出十二種類型的學習曲線，其目的在探討四種分析模式的估計成效及模式的估計會受到哪些因素所影響。

圖 9 類型 EH、E、EL 學習曲線圖



### 三、實徵資料

本研究採用教育部推動之學習平臺——因材網（教育部，2017）中的動態評量（dynamic assessment）做為實徵資料。動態評量是一種將教學嵌入評量的教學方法，常見的形式為先測驗（test），接著中介教學（intervention），再進行測驗（retest），如圖 10 所示。有許多研究表示，動態評量可使學生在評量期間獲得即時且有意義的反饋，讓學生可以進行自我反思，並提高其學習動機與成效（Jitendra et al., 1994; Marriott, 2009; Wang, 2011）。電腦化的動態評量則可降低教師在實施動態評量時，所花費的時間及教學人力上的負擔。在因材網中的動態評量為單選題之測驗形式，每題皆為四選一的選擇題，如圖 11 所示。



測驗學刊，第 71 輯第 3 期

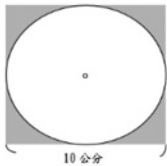
圖 10 動態評量常見形式



圖 11 因材網的動態評量

(第1次作答)

Q1 下圖為一個黃色色紙，剪下一個直徑為 10 公分的圓面積，請問剩下的黃色色紙面積是多少平方公分？



10 公分

---


A

☒ 100

☐ 78.5

☐ 21.5

☐ 67

 作答完畢

學生在第一次作答錯誤時，系統會給予一般性的提示（如圖 12 所示），如果第二次作答仍答錯則會給予較關鍵性的提示（如圖 13 所示），第三次仍然答錯則會進行詳解教學（如圖 14 所示）。

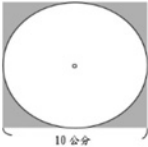
#### 四、評估指標

相關模式之評估多以 AIC 及 BIC 做為模式適配度指標（Cen et al., 2006; DataShop, n.d.; Liu & Koedinger, 2017），故本研究將計算估計誤差，以判斷模式估計精準度，並以 AIC（Akaike information criterion; Akaike, 1998）及 BIC（Bayesian information criterion; Schwarz, 1978）做為模式適配度指標。

圖 12 第一次作答錯誤給予一般性提示

(第2次作答)

**Q1** 下圖為一個黃色色紙，剪下一個直徑為 10 公分的圓面積，請問剩下的黃色色紙面積是多少平方公分？



10 公分

---

**提示**

不對喔!!題目看清楚了嗎?是要算剩下的部分喔?想一想,再算一次吧!!!

**A**

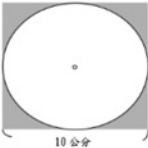
☐ 100  
☒ 78.5  
☐ 21.5  
☐ 67

⏪ 作答完畢

圖 13 第二次作答錯誤給予關鍵性提示

(第3次作答)

**Q1** 下圖為一個黃色色紙，剪下一個直徑為 10 公分的圓面積，請問剩下的黃色色紙面積是多少平方公分？



10 公分

---

**提示**

答錯了!!本題要先算出正方形面積，再用正方形面積減掉內部的圓面積，即可求出答案。再算一次吧。

**A**

☐ 100  
☐ 78.5  
☐ 21.5  
☒ 67

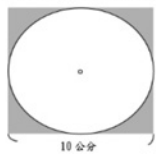
⏪ 作答完畢

測驗學刊，第 71 輯第 3 期

圖 14 第三次作答錯誤給予詳解教學

(第4次作答)

**Q1**下圖為一個黃色色紙，剪下一個直徑為 10 公分的圓面積，請問剩下的黃色色紙面積是多少平方公分？



10 公分

---

**正確解法**

正方形面積：10×10=100 (平方公分)  
 內部減掉的圓面積：5×5×3.14=78.5  
 100-78.5=21.5 (平方公分)，所以正確答案為選項③

**A**

☐ 100  
☐ 78.5  
☐ 21.5  
☐ 67

作答完畢  
 進入下一題

### (一) 誤差

計算誤差的公式如下：

$$E = \frac{\sum_{t=1}^T |R_t - P_t|}{T} \quad (11)$$

其中， $E$  為平均誤差， $R_t$  為在第  $t$  次測驗的實際錯誤率， $P_t$  為在第  $t$  次測驗的預測錯誤率， $T$  為測驗次數。

### (二) -2LL

-2LL (log likelihood) 的公式如下：

$$-2LL = -2\log\left[\prod_{i=1}^n L(X_i, \hat{\theta}_i, \hat{\xi})\right] \quad (12)$$

郭伯臣 林佳樺 葉珮玉 楊智為

學習曲線分析於電腦化動態評量之應用

其中， $n$  表示樣本數， $X_i$  表示作答反應， $\theta_i$  表示學生的能力， $\xi$  表示估計的試題參數。 $-2LL$  值愈小，表示模式與資料的適配度愈佳。在本研究中，僅做為 AIC 及 BIC 評估指標之計算用。

### (三) AIC

AIC 的公式如下：

$$AIC = -2LL + 2P \quad (13)$$

其中， $P$  為估計的模式參數個數。AIC 值愈小，表示模式與資料的適配度愈佳，AIC 值為  $-2LL$  值加上懲罰值  $2P$ ，當模式愈複雜，估計的參數愈多，懲罰就愈重。

### (四) BIC

BIC 的公式如下：

$$BIC = -2LL + P \log(N) \quad (14)$$

其中， $P$  表示模式要估計的參數個數， $N$  表示樣本數。BIC 值愈小，表示模式與資料的適配度愈佳，BIC 值為  $-2LL$  值加上懲罰值  $p \log(N)$ ，懲罰值同時考慮到模式參數個數與樣本數。

## 肆、研究結果

### 一、模擬實驗之研究結果

表 3 為理想學習曲線之估計誤差，其誤差結果為每個類型模擬 10 次的平均誤差值。

測驗學刊，第 71 輯第 3 期

表 3 理想學習曲線之估計誤差

| 類型 | AFM  | iAFM | BKT   | iBKT |
|----|------|------|-------|------|
| A  | 1.86 | 1.84 | 0.86  | 3.99 |
| AL | 1.31 | 1.12 | 0.68  | 3.07 |
| B  | 3.28 | 3.25 | 3.96  | 3.65 |
| BL | 0.58 | 0.46 | 2.40  | 3.33 |
| C  | 3.46 | 3.59 | 10.68 | 3.94 |
| CL | 3.52 | 3.65 | 6.48  | 4.07 |

圖 15～圖 20 為類型 A、AL、B、BL、C 及 CL 的學習曲線圖，皆為理想學習曲線圖。四種模式的預測趨勢皆具有一致性，須從誤差來判斷何種模式估計較為精準。

類型 A 及 AL 屬於學生學習率較高的資料型態（斜率較大），由圖 15 與圖 16 可知，學生在經過第一次教學介入後，錯誤率有大幅下降的趨勢，從原本將近 90% 及 80% 降至 10% 左右，其中估計誤差最小的模式皆為 BKT 模式。

在類型 B 及 BL 中，學生的學習率並沒有類型 A 及 AL 高，斜率較為趨緩，至少需經過兩次的教學介入後，學生的錯誤率才會降到 10% 左右，其中估計誤差最小的模式皆為 iAFM。

在類型 C 及 CL 中，學生的學習率偏低（斜率最小），至少需要經過三至四次的教學介入後，學生的錯誤率才會降到 10% 左右，其中估計誤差最小的模式皆為 AFM。

圖 15 類型 A 學習曲線圖

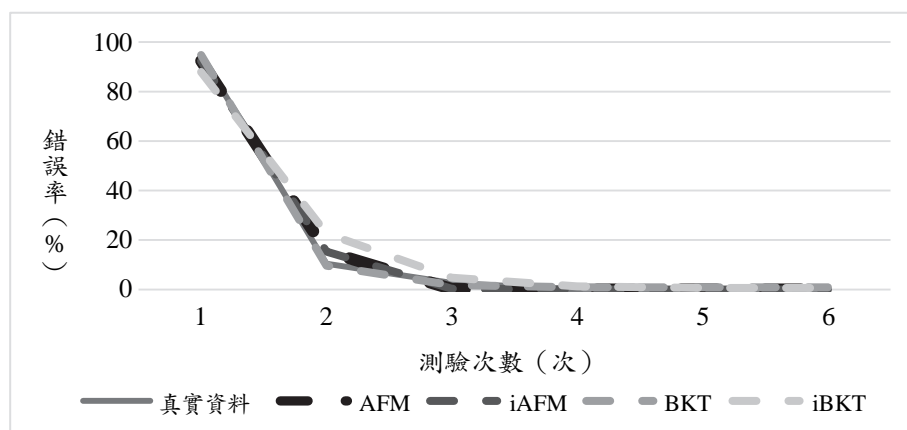


圖 16 類型 AL 學習曲線圖

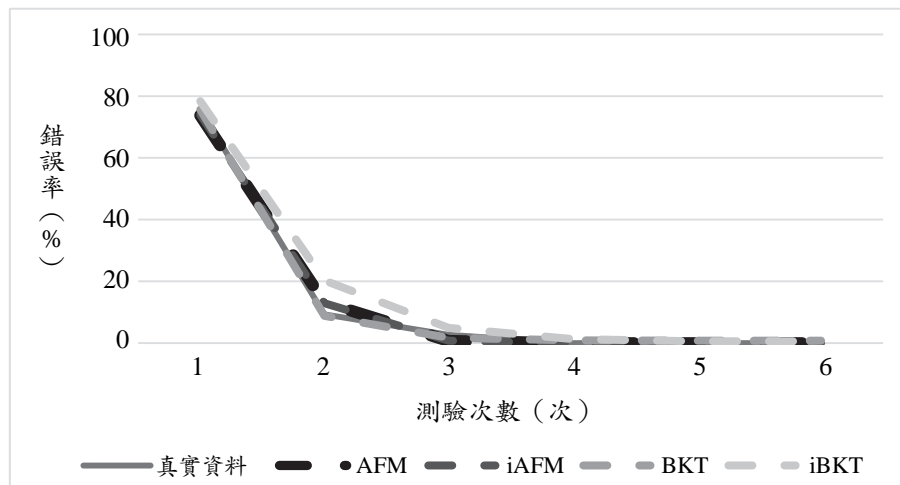
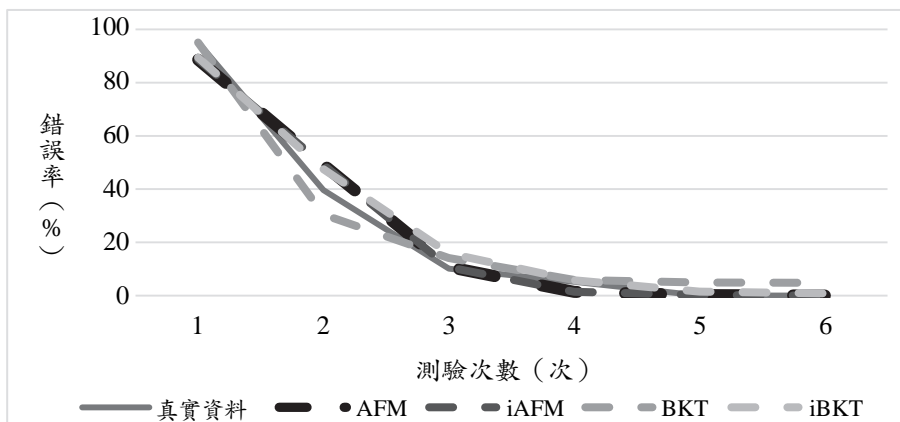


圖 17 類型 B 學習曲線圖



由實驗結果可知，學生的學習率較高時，BKT 估計較為精準；學習率較為趨緩時，iAFM 估計較為精準；學習率較低時，為 AFM 估計較為精準。模式的估計成效不會因為學生不同的起始能力而改變，但會受到學生不同的學習率而影響。

測驗學刊，第 71 輯第 3 期

圖 18 類型 BL 學習曲線圖

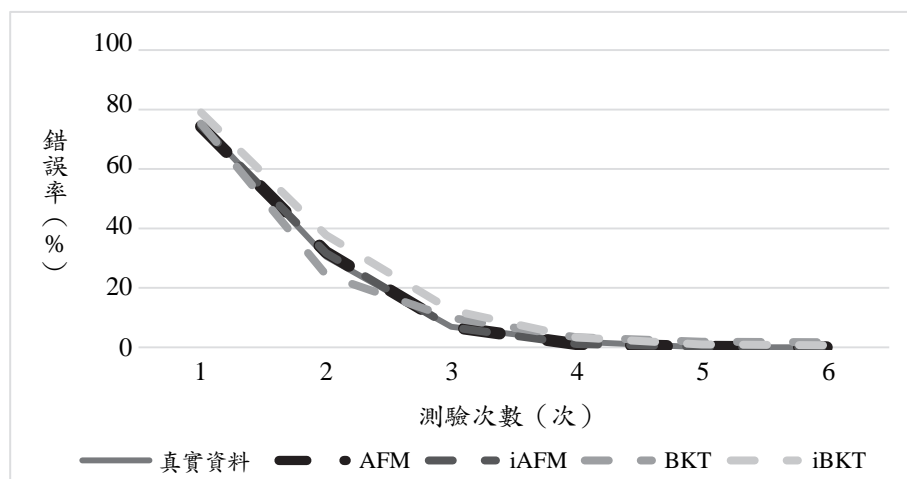


圖 19 類型 C 學習曲線圖

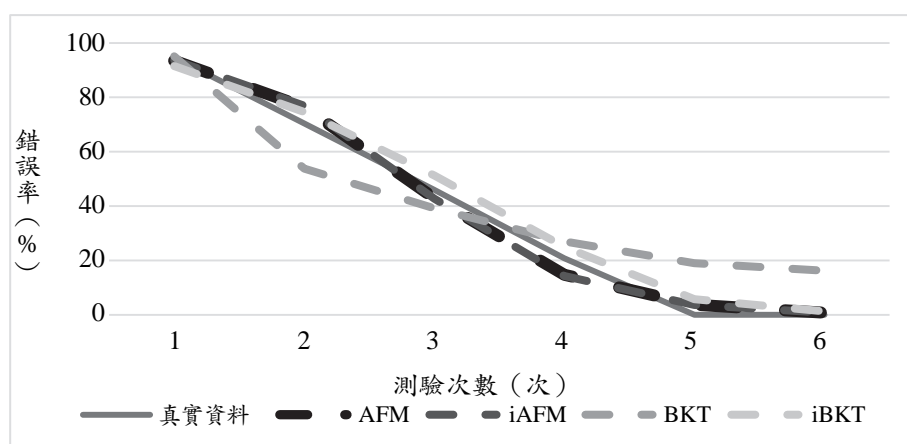




圖 20 類型 CL 學習曲線圖

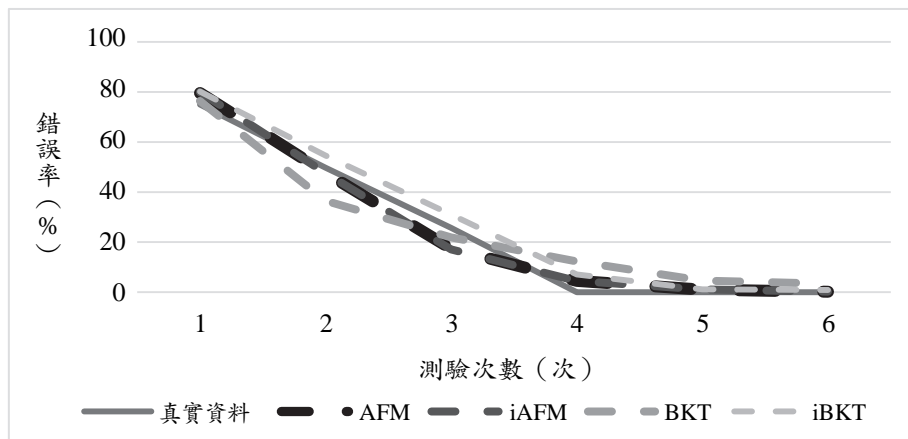


圖 21～圖 26 分別為類型 DH、D、DL、EH、E 及 EL 的學習曲線圖。DH、D 及 DL 三種類型為學習平緩學習曲線，依據學生不同的起始能力分成三類；EH、E 及 EL 屬於學習無效學習曲線，同樣也依據學生不同的起始能力分成三類，以便探討不同學習型態、學生起始能力、學生學習率對於模式估計的影響。

由圖 21～圖 23 及表 4 可知，在預測學習平緩學習曲線時，AFM 預測最精準；在類型 D 時，iAFM 的預測誤差與 AFM 相同。AFM 及 iAFM 的估計成效比 BKT 及 iBKT 好，且估計的穩定性也較高，不會因為學生的學習率改變而造成估計的精準度有明顯差異。

由圖 24～圖 26 及表 5 可知，在預測學習無效學習曲線時，AFM 的估計精準度較佳。

測驗學刊，第 71 輯第 3 期

圖 21 類型 DH 學習曲線圖

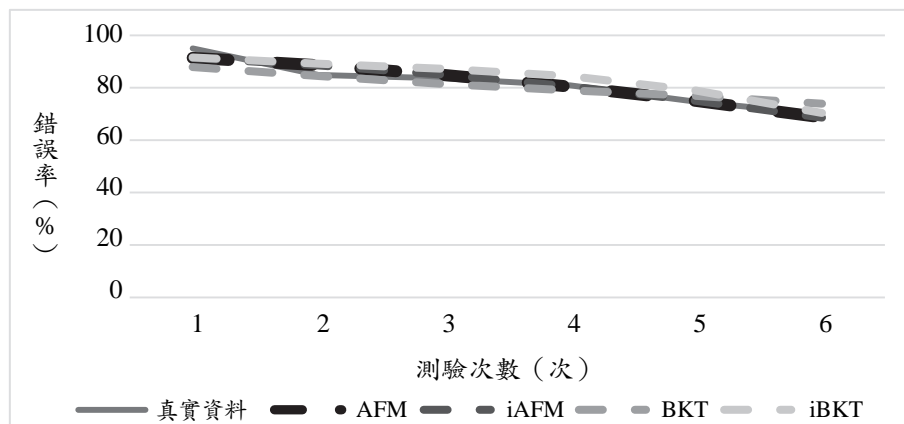
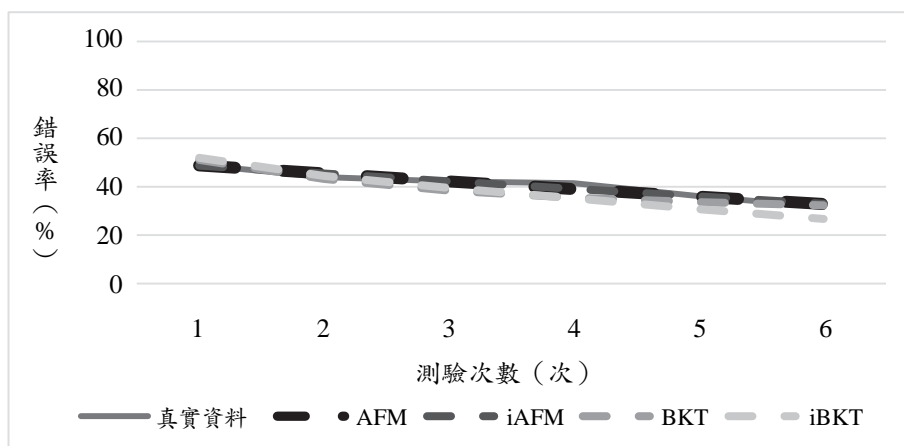


圖 22 類型 D 學習曲線圖



郭伯臣 林佳樺 葉珮玉 楊智為

學習曲線分析於電腦化動態評量之應用

圖 23 類型 DL 學習曲線圖

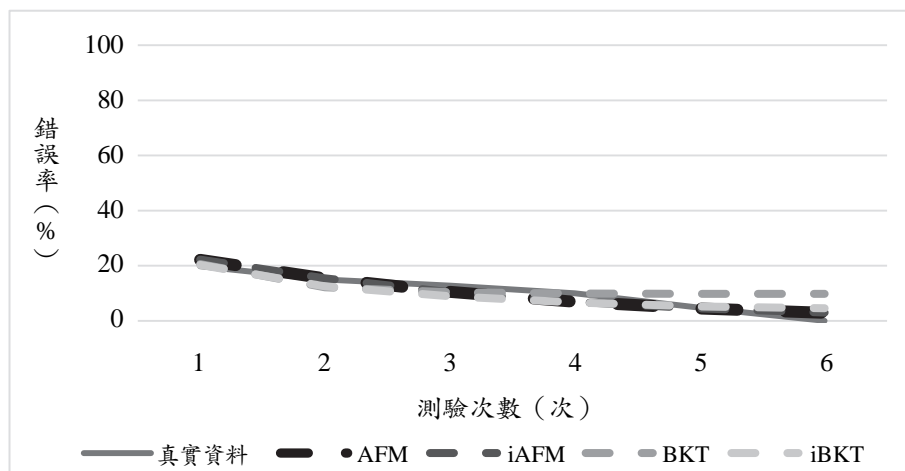
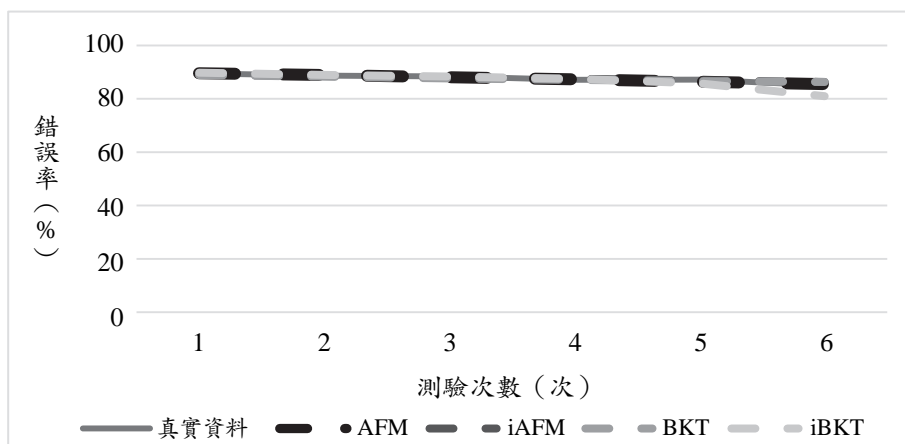


圖 24 類型 EH 學習曲線圖



測驗學刊，第 71 輯第 3 期

圖 25 類型 E 學習曲線圖

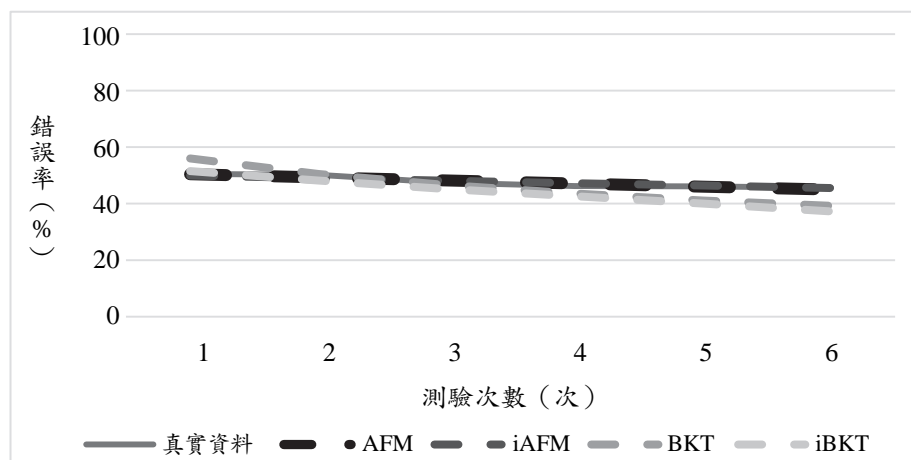


圖 26 類型 EL 學習曲線圖

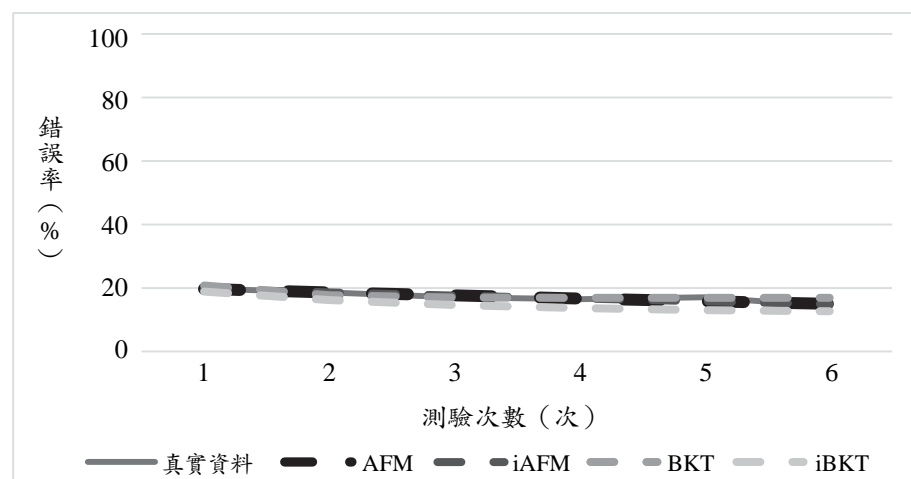


表 4 學習平緩學習曲線之估計誤差

| 類型 | AFM  | iAFM | BKT  | iBKT |
|----|------|------|------|------|
| DH | 1.78 | 1.79 | 2.82 | 3.31 |
| D  | 0.98 | 0.98 | 2.74 | 3.54 |
| DL | 1.90 | 1.92 | 3.53 | 2.53 |

表 5 學習無效學習曲線之估計誤差

| 類型 | AFM  | iAFM | BKT  | iBKT |
|----|------|------|------|------|
| EH | 0.42 | 0.46 | 0.79 | 1.13 |
| E  | 0.59 | 0.68 | 3.86 | 4.00 |
| EL | 0.48 | 0.52 | 1.05 | 2.39 |

結果顯示，不管是哪一種學習曲線，其模式估計成效皆為 AFM 及 iAFM 優於 BKT 及 iBKT，只有在類型 A 及類型 AL 為 BKT 預測的效果最佳。同樣的，AFM 及 iAFM 預測誤差並無明顯差異，因此從模式適配度來進行比較，如表 6 所示。

由表 6 的模式適配次數表可知，除了在類型 A 及 AL 的 AIC 為 BKT 較小之外，其他皆為 iAFM 的適配度較佳，明顯優於 AFM。在考慮精準度和適配度的情況下，會將 iAFM 列為選擇模式的第一順位。此外，iAFM 還能考慮到學生的個人參數，是 AFM 所沒有的。

表 6 各類型模式的適配次數表

| 類型 | AIC |      |     |      | BIC 值 |      |     |      |
|----|-----|------|-----|------|-------|------|-----|------|
|    | AFM | iAFM | BKT | iBKT | AFM   | iAFM | BKT | iBKT |
| A  | 0   | 0    | 10  | 0    | 0     | 5    | 5   | 0    |
| AL | 0   | 3    | 7   | 0    | 0     | 7    | 3   | 0    |
| B  | 0   | 10   | 0   | 0    | 0     | 10   | 0   | 0    |
| BL | 0   | 10   | 0   | 0    | 0     | 10   | 0   | 0    |
| C  | 0   | 10   | 0   | 0    | 0     | 10   | 0   | 0    |
| CL | 0   | 10   | 0   | 0    | 0     | 10   | 0   | 0    |
| DH | 0   | 10   | 0   | 0    | 0     | 10   | 0   | 0    |
| D  | 0   | 10   | 0   | 0    | 0     | 10   | 0   | 0    |
| DL | 0   | 10   | 0   | 0    | 0     | 10   | 0   | 0    |
| EH | 0   | 10   | 0   | 0    | 0     | 10   | 0   | 0    |
| E  | 0   | 10   | 0   | 0    | 0     | 10   | 0   | 0    |
| EL | 0   | 10   | 0   | 0    | 0     | 10   | 0   | 0    |

從模擬資料的研究結果顯示，不管在哪一種類型的學習曲線預測上，AFM 及 iAFM 的預測精準度皆高於 BKT 及 iBKT，其模式的穩定性也較佳，不會因為其他因素而導致估計誤差變大。學生不同的起點行為並不會對模式

的表現有影響，但在不同的學習曲線中，學生的學習率會影響模式的表現。在理想學習曲線中，學生不同的學習率會影響估計的表現，在類型 A（學習率高）時為 BKT 模式表現較佳；在類型 B 及 C 時，分別為 iAFM 及 AFM 的預測較好，但在學習平緩及學習無效的學習曲線上，學生的學習率對於模式的表現沒有太大影響，皆為 AFM 模式的表現最好。

## 二、實徵資料之研究結果

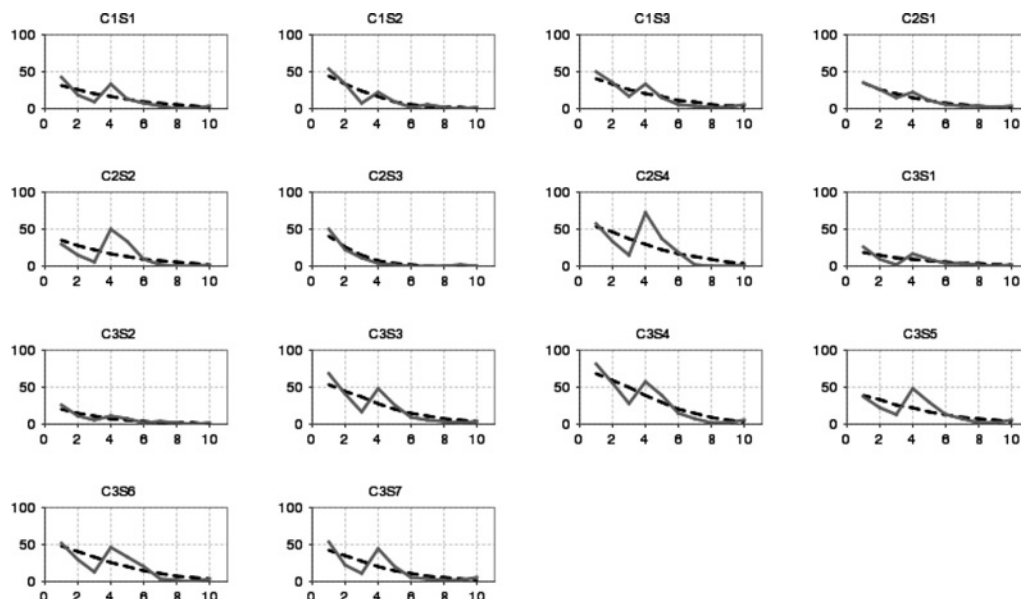
實徵資料採用因材網六年級數學：圓面積單元的動態評量作答資料，其資料集包含 900 位學生、72,287 筆作答資料，該單元有 3 個能力指標（In），並由學科專家拆解成 14 個子技能（Skill），每個技能各有 2 題動態評量。在測驗次數上，會依據每位學生的作答情形不同而有所差異，透過 AFM、iAFM、BKT 及 iBKT 評估及預測學生的學習情況，除了比較四個模式之模式適配度外，本研究亦探討不同 KC 模式的模式適配度，並做概念拆解之示範，以找出最符合學生真實學習情形之模式。

由表 7 可知，四種估計模式在拆解成 14 個子技能時，模式適配度較佳，其中 iAFM 的 AIC 及 BIC 與實徵資料的適配度最高。圖 27 為 14 個子技能之學習曲線圖，x 軸為學生的練習次數，y 軸為學生的學習表現（錯誤率），實線為學生的真實資料，虛線為 iAFM 預測的曲線，14 個子技能之學習曲線圖都屬於理想學習曲線，代表動態評量中的教學介入（提示）對學生而言是有幫助的。但是，大部分的子技能在第 4 次練習時都出現尖點，在重新檢視動態評量之試題後發現，相同的子技能在不同試題上因難度或所需的先備知識

表 7 不同估計模式與 KC 模式之 AIC 及 BIC

| 估計模式 | KC 模式 | 概念數 | AIC              | BIC               |
|------|-------|-----|------------------|-------------------|
| AFM  | In    | 3   | 37,206.72        | 45,527.93         |
|      | Skill | 14  | <b>25,097.19</b> | <b>33,620.46</b>  |
| iAFM | In    | 3   | 36,432.74        | 36,506.22         |
|      | Skill | 14  | <b>23,810.76</b> | <b>23,884.24</b>  |
| BKT  | In    | 3   | 38,597.98        | 38,708.20         |
|      | Skill | 14  | <b>28,688.50</b> | <b>29,202.83</b>  |
| iBKT | In    | 3   | 65,785.59        | 334,848.22        |
|      | Skill | 14  | <b>33,931.69</b> | <b>132,490.33</b> |

圖 27 14 個子技能的學習曲線圖



有所差異，故在換題時造成尖點出現。

C2S2 的子技能為「能算出複合圖形的周長」，但該子技能的兩題圖形分別為「兩個正方形」與「正方形內嵌半圓」，第 2 題多涵蓋了半圓周長，故在第 1 題動態評量（含 3 次練習機會）後，進入第 2 題（第 4 次練習）時出現尖點。C3S6 的子技能為「能計算扇形的面積」，但該子技能的兩題角度分別為 180 度與 120 度。C3S7 的子技能為「能運用圓面積公式解決問題（計算陰影區域之面積）」，但該子技能的兩題圖形分別是「正方形與圓形」、「正方形與三角形」。這三個子技能之動態評量試題皆非複本題，因此本研究將這 3 種子技能拆解成 6 種概念，並生成一個新的 KC 模式。

由表 8 可知，在不同的估計模式下，拆解後新的 KC 模式之模式適配度皆優於表 7 中 KC 模式的 Skill，其中 iAFM 之模式適配度最佳。拆解後的 C2S2、C3S6 與 C3S7 的學習曲線如圖 28 所示。由圖可知，拆解後的概念呈現更為平滑的理想學習曲線。



測驗學刊，第 71 輯第 3 期

表 8 新的 KC 模式（概念數 = 17）之 AIC 及 BIC

| 估計模式 | AIC      | BIC       |
|------|----------|-----------|
| AFM  | 24420.16 | 32998.54  |
| iAFM | 23072.60 | 23146.08  |
| BKT  | 26460.70 | 27085.25  |
| iBKT | 43646.39 | 171858.75 |

圖 28 拆解後的學習曲線

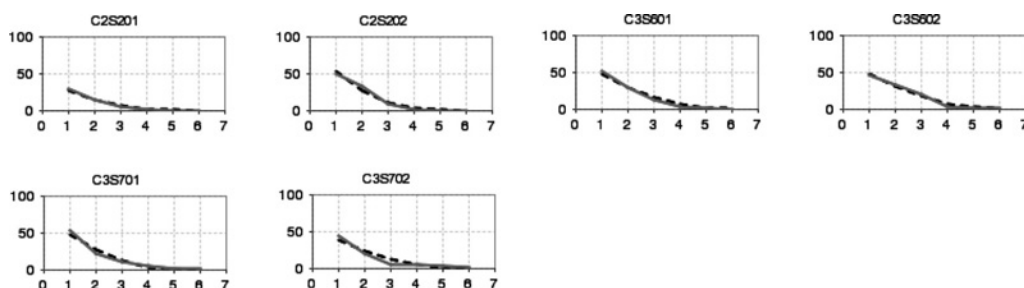


圖 29～圖 31 為三位學生的學習曲線圖，學生 1 的作答反應為 {0,0,0,1,1,1}；學生 2 的作答反應為 {1,1,1,1,1,1}；學生 3 的作答反應為 {0,0,0,0,0,1}，實線為學生真實作答的情形，短虛線為 AFM 所估計，長虛線為 iAFM 估計的曲線。由圖可以看出，iAFM 因為考量學生的個人學習率，估計結果更符合學生實際作答的情形，因此本研究會建議選用 iAFM 做為估計模式。

另外，有一些學習曲線在第 4 次練習時也會出現尖點（如圖 27 所示），例如：C1S1、C2S4 與 C3S5，這些子技能的動態評量皆為複本題，表示前 3 次練習的錯誤率會隨著練習次數增加而下降，此可能是因為在選擇題中逐次排除錯誤選項，而非動態評量的提示或教學所帶來的效果。

圖 29 學生 1 之學習曲線圖

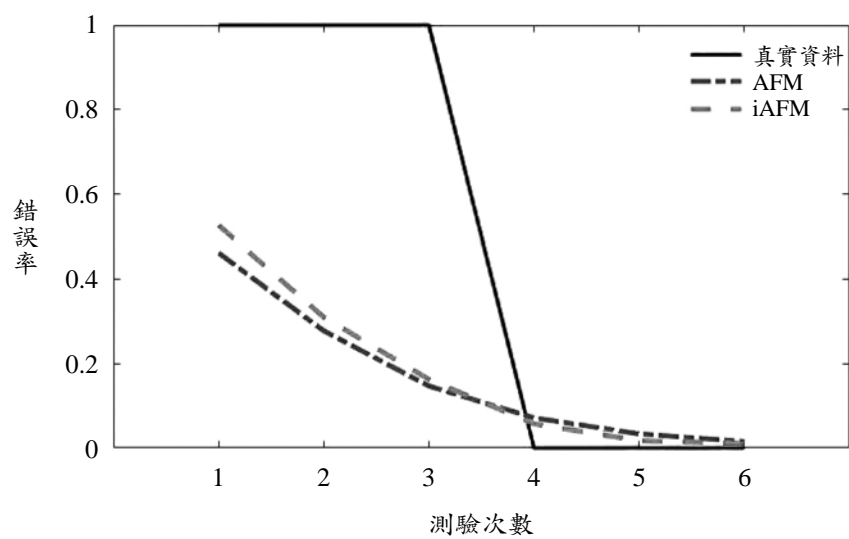
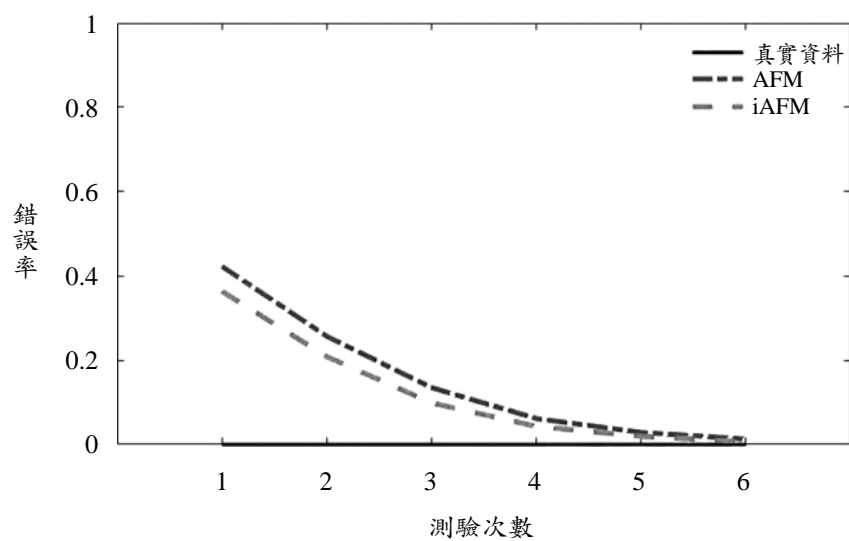
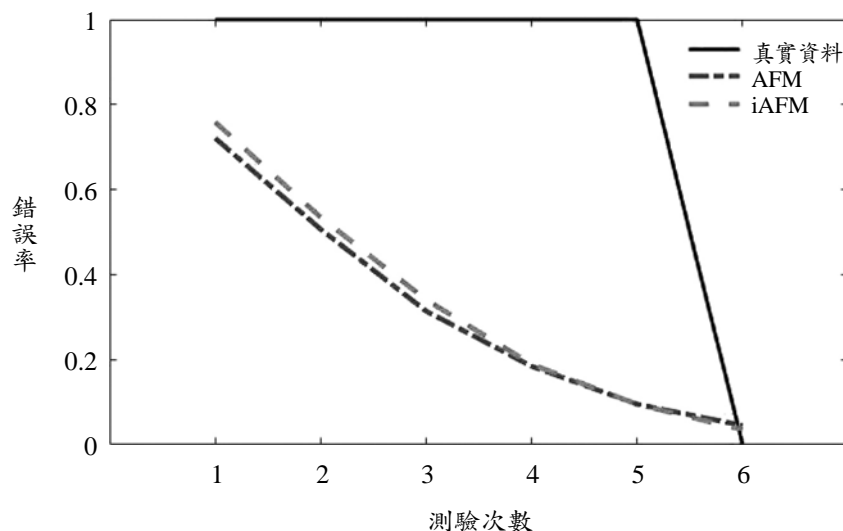


圖 30 學生 2 之學習曲線圖



測驗學刊，第 71 輯第 3 期

圖 31 學生 3 之學習曲線圖



## 伍、結論與建議

### 一、模式估計之精準度

模擬研究的結果顯示，四種模式的估計趨勢皆具有一致性，AFM 及 iAFM 的估計成效明顯比 BKT 及 iBKT 佳，其估計的誤差值較小，穩定度相對也較高，並不會因為學習曲線類型的不同，導致估計誤差有太大差異。而 AFM 及 iAFM 兩者估計的誤差值相近，但 iAFM 的模式適配度較 AFM 高，此可能是因為 iAFM 能適性調整學生的學習率，使得模式能夠更準確的描述學生的學習過程和表現，並能更好的解釋學生之間的個別差異，而提高了模式的預測能力。因此，透過 iAFM 可使教師藉由個人學習率給予學生適性的反饋，並能針對學生不同的需求，給予不同的教學，提高學生的學習成效。

### 二、學生的起點行為與學習率對模式估計的影響

研究結果顯示，不同模式不會因為學生的起始能力高低而有所不同。在不同類型的學習曲線中，學生的學習率對於模式估計之表現有所不同。在理想學習曲線中，學生的學習率會影響模式估計的成效，在類型 A 時為 BKT 模

式表現的最優異，這可能是因為 BKT 在建模的過程中，將學生的知識狀態分為二元變量：已習得（learned）和未習得（not learned），當學生在一次學習後的錯誤率明顯降低時，模式就更能夠確定學生已經掌握了特定的知識或技能，因此更可能將其歸為「已習得」的狀態，進而提供更準確的估計。而在估計類型 B 及 C 的學習曲線中，分別為 iAFM 及 AFM 的表現較佳，但在學習平緩及學習無效的學習曲線上，學生的學習率對於模式的表現沒有太大影響，皆為 AFM 模式的表現最好，其估計的精準度誤差值也都相當小。

綜合以上所述，模式估計表現不受學生起點行為的影響，在不同學習曲線類型的估計中，除了在一次學習後的錯誤率大幅下降時，BKT 模式表現較佳，其餘皆為 AFM 及 iAFM 的估計精準度優於 BKT 及 iBKT，且 AFM 及 iAFM 的估計精準度差不多，而在模式適配度檢驗下，無論是模擬研究或是實證研究，皆以 iAFM 模式適配度較優於 AFM，故在只能選用其中一種模式的情況下，建議研究者可選用 iAFM，未來並可結合 BKT 及 iAFM 的優勢發展新的模式。

### 三、學習曲線分析之應用

透過學習曲線圖，教師可以觀察到是否存在複雜概念，再進一步拆解後，可由概念難易度調整教學順序，就能更符合學生的學習狀況。研究者或教師可以透過學習曲線分析，了解學生的程度，當學習曲線呈現高平時，可以調整教學的內容或增加學生練習的次數；當學習曲線為低平時，可以減少學生練習的次數，避免學生花過多的時間練習已經掌握的技能；當學習曲線為學習無效時，可以調整教學教材。教師也可以利用學習曲線預測學生未來的學習表現，像是可以利用平常在校的成績來預測學生在升學考試時的表現。在學習平臺應用上，可以從過往學生所建構好的學習曲線來預測新進學生的能力，給予其適當的學習規劃與建議。

透過實徵研究可知，學習平臺提供的提示或教學可提升學生的學習成效，但選擇題型的動態評量可能會因為猜測或逐次排除錯誤選項，而導致誤判學生真實的學習情形，因此建議未來動態評量的題型可朝建構反應題發展，並藉由人工智慧技術加入提示或教學。

## 謝誌

本研究使用教育部因材網六年級數學「圓面積」單元進行實徵研究，特此感謝教育部「推動中小學數位學習精進方案」及「適性教學全國推動計畫」提供的寶貴資料與支持。使用本資料發表之學術論文僅代表作者本人的觀點，不代表教育部的觀點、看法與立場，相關責任作者自負。另外，感謝審查委員對於本文提供的寶貴修改建議，及編輯委員的細心校對。

## 參考文獻

### 中文部分

林佳樺、陳俊華、郭伯臣（2018 年 10 月）。**BKT 與 AFM 於因材網學習資料分析比較**。發表於第十三屆海峽兩岸心理與教育測驗學術研討會暨中國測驗學會年會，南投。

[Lin, C.-H., Chen, C.-H., & Kuo, B.-C. (2018). *Comparison of BKT and AFM in analyzing Taiwan Adaptive Learning Platform data*. The 13th Cross-Strait Conference on Educational and Psychological Testing Annual Meeting of Chinese Association of Psychological Testing, Nantou, Taiwan.]

教育部（2017）。**適性測驗與學習平台：因材網**。（<https://adl.edu.tw/>）

[Ministry of Education. (2017). *Adaptive Assessment and Learning Platform: Taiwan Adaptive Learning Platform*. (<https://adl.edu.tw/>)]

葉珮玉（2020）。**不同學習曲線分析模式之比較研究**〔未出版之碩士論文〕。國立臺中教育大學。

[Ye, P.-Y. (2020). *A comparative study of learning curve analysis models* [Unpublished master's thesis]. National Taichung University of Education.]

葉珮玉、林佳樺、郭伯臣、陳俊華（2019 年 5 月）。**國小數學電腦化動態評量之學習分析研究**。發表於 2019 全球華人教育資訊與評估研討會，國立臺中教育大學。

[Ye, P.-Y., Lin, C.-H., Kuo, B.-C., & Chen, C.-H. (2019). *Study of learning analysis in computerized dynamic assessment of mathematics for elementary school*. 2019 Global Chinese Conference on Educational Information and Assessment, National Taichung University of Education.]

### 英文部分

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected papers of Hirotugu Akaike* (pp. 199-213). Springer. [https://doi.org/10.1007/978-1-4612-1694-0\\_15](https://doi.org/10.1007/978-1-4612-1694-0_15)

Baker, R. S. J. d., Corbett, A. T., & Aleven, V. (2008a). *Improving contextual models of*

- guessing and slipping with a truncated training set*. <https://learninganalytics.upenn.edu/ryanbaker/EDMGuessSlip2008Q.pdf>
- Baker, R. S. J. d., Corbett, A. T., & Aleven, V. (2008b). More accurate student modeling through contextual estimation of slip and guess probabilities in bayesian knowledge tracing. In B. P. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent tutoring systems. ITS 2008. Lecture Notes in Computer Science* (vol 5091). Springer. [https://doi.org/10.1007/978-3-540-69132-7\\_44](https://doi.org/10.1007/978-3-540-69132-7_44)
- Bulut, O., Shin, J., Yildirim-Erbasli, S. N., Gorgun, G., & Pardos, Z. A. (2023). An introduction to Bayesian knowledge tracing with pyBKT. *Psych*, 5(3), 770-786.
- Cen, H. (2009). *Generalized learning factors analysis: Improving cognitive models with machine learning*. Carnegie Mellon University.
- Cen, H., Koedinger, K., & Junker, B. (2006, June). Learning factors analysis: A general method for cognitive model evaluation and improvement. In M. Ikeda, K. D. Ashley, & T. W. Chan (Eds.), *Intelligent Tutoring Systems. ITS 2006. Lecture Notes in Computer Science* (vol 4053). Springer. [https://doi.org/10.1007/11774303\\_17](https://doi.org/10.1007/11774303_17)
- Cen, H., Koedinger, K., & Junker, B. (2008). Comparing two IRT models for conjunctive skills. In B. P. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent Tutoring Systems. ITS 2008. Lecture Notes in Computer Science* (vol 5091). Springer. [https://doi.org/10.1007/978-3-540-69132-7\\_111](https://doi.org/10.1007/978-3-540-69132-7_111)
- Cloudspurs (2019). *Bayesian Knowledge Tracing*. <https://github.com/cloudspurs/bkt>
- Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-adapted Interaction*, 4(4), 253-278. <https://doi.org/10.1007/BF01099821>
- DataShop (n.d.). *The PSLC DataShop*. <https://pslcdatashop.web.cmu.edu/>
- David, Y. B., Segal, A., & Gal, Y. A. (2016, April). Sequencing educational content in classrooms using Bayesian knowledge tracing. In *LAK 2016 Conference Proceedings, 6th International Learning Analytics and Knowledge Conference - Enhancing Impact: Convergence of Communities for Grounding, Implementation, and Validation* (pp. 354-363). (ACM International Conference Proceeding Series; Vol. 25-29-April-2016). Association for Computing Machinery. <https://doi.org/10.1145/1235>
- Durand, G., Goutte, C., Belacel, N., Bouslimani, Y., & Leger, S. (2017). Review, computation and application of the Additive Factor Mode l (AFM). *Nat. Res. Council Canada, Winnipeg, MB, Canada, Tech. Rep, 23002483*. <https://doi.org/10.4224/23002483>
- Goldin, I., Pavlik Jr., P. I., & Ritter, S. (2016). Discovering domain models in learning curve data. *Design Recommendations for Intelligent Tutoring Systems: Volume 4-Domain*

- Modeling*, 4, 115-126.
- Jitendra, A. K., Kameenui, E. J., & Carnine, D. W. (1994). An exploratory evaluation of dynamic assessment and the role of basals on comprehension of mathematical operations. *Education and Treatment of Children*, 139-162. <https://www.jstor.org/stable/42899350>
- Klingler, S., Kaser, T., Solenthaler, B., & Gross, M. (2015). On the performance characteristics of latent-factor and knowledge tracing models. *International Educational Data Mining Society*.
- Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representation on quantitative reasoning. *Journal of the Learning Sciences*, 13(2), 129-164. [https://doi.org/10.1207/s15327809jls1302\\_1](https://doi.org/10.1207/s15327809jls1302_1)
- Koedinger, K. R., Carvalho, P. F., Liu, R., & McLaughlin, E. A. (2023). An astonishing regularity in student learning rate. *Proceedings of the National Academy of Sciences*, 120(13), e2221311120.
- Koedinger, K. R., Stamper, J. C., McLaughlin, E. A., & Nixon, T. (2013). Using data-driven discovery of better student models to improve student learning. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Artificial intelligence in education. AIED 2013. Lecture Notes in Computer Science* (vol 7926). Springer. [https://doi.org/10.1007/978-3-642-39112-5\\_43](https://doi.org/10.1007/978-3-642-39112-5_43)
- Koedinger, K. R., Yudelson, M. V., & Pavlik Jr., P. I. (2016). *Testing theories of transfer using error rate learning curves*. *Topics in Cognitive Science*, 8(3), 589-609. <https://doi.org/10.1111/tops.12208>
- Learnsphere (2018). *Learnsphere*. <http://learnsphere.org/>
- Liu, R., & Koedinger, K. R. (2015). Variations in learning rate: Student classification based on systematic residual error patterns across practice opportunities (ED-560874). *International Educational Data Mining Society*.
- Liu, R., & Koedinger, K. R. (2017). Towards reliable and valid measurement of individualized student parameters. *International Educational Data Mining Society*.
- Marriott, P. (2009). Students' evaluation of the use of online summative assessment on an undergraduate financial accounting module. *British Journal of Educational Technology*, 40(2), 237-254. <https://doi.org/10.1111/j.1467-8535.2008.00924.x>
- Nathan, M. J., & Koedinger, K. R. (2000a). An investigation of teachers' beliefs of students' algebra development. *Cognition and Instruction*, 18(2), 209-237. [https://doi.org/10.1207/S1532690XCI1802\\_03](https://doi.org/10.1207/S1532690XCI1802_03)
- Nathan, M. J., & Koedinger, K. R. (2000b). Teachers' and researchers' beliefs about the de-



- velopment of algebraic reasoning. *Journal for Research in Mathematics Education*, 31(2), 168-190. <https://doi.org/10.2307/749750>
- Nathan, M. J., & Petrosino, A. (2003). Expert blind spot among preservice teachers. *American Educational Research Journal*, 40(4), 905-928. <https://doi.org/10.3102/00028312040004905>
- Nathan, M. J., Long, S. D., & Alibali, M. W. (2002). The symbol precedence view of mathematical development: A corpus analysis of the rhetorical structure of algebra textbooks. *Discourse Processes*, 33(1), 1-21. [https://doi.org/10.1207/S15326950DP3301\\_01](https://doi.org/10.1207/S15326950DP3301_01)
- Newell, A., & Rosenbloom, P. S. (1981). *Mechanisms of skill acquisition and the law of practice*. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Lawrence Erlbaum Associates
- Pardos, Z. A., & Heffernan, N. T. (2010, June). Modeling individualization in a bayesian networks implementation of knowledge tracing. In P. De Bra, A. Kobsa, & D. Chin (Eds.), *User Modeling, Adaptation, and Personalization. UMAP 2010. Lecture Notes in Computer Science* (vol 6075). Springer. [https://doi.org/10.1007/978-3-642-13470-8\\_24](https://doi.org/10.1007/978-3-642-13470-8_24)
- Qiu, Y., Qi, Y., Lu, H., Pardos, Z. A., & Heffernan, N. T. (2011, July). Does time matter? Modeling the effect of time with Bayesian knowledge tracing. In *Proceedings of the 4th International Conference on Educational Data Mining* (pp. 139-148).
- Rachatasumrit, N., & Koedinger, K. R. (2021). Toward improving student model estimates through assistance scores in principle and in practice. *International Educational Data Mining Society*.
- Rivers, K., Harpstead, E., & Koedinger, K. R. (2016). Learning curve analysis for programming: Which concepts do students struggle with? In *Proceedings of the 2016 ACM Conference on International Computing Education Research* (pp. 143-151). <https://doi.org/10.1145/2960310.2960333>
- Sao Pedro, M., Baker, R., & Gobert, J. (2013, July). Incorporating scaffolding and tutor context into Bayesian knowledge tracing to predict inquiry skill acquisition. [https://www.educationaldatamining.org/EDM2013/papers/rn\\_paper\\_27.pdf](https://www.educationaldatamining.org/EDM2013/papers/rn_paper_27.pdf)
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461-464.
- Wang, S., Han, Y., Wu, W., & Hu, Z. (2017, April). Modeling student learning outcomes in studying programming language course. In *2017 Seventh International Conference on Information Science and Technology (ICIST)* (pp. 263-270). IEEE.

- Wang, T. H. (2011). Implementation of Web-based dynamic assessment in facilitating junior high school students to learn mathematics. *Computers & Education*, 56(4), 1062-1071. <https://doi.org/10.1016/j.compedu.2010.09.014>
- Yan, D.-T., Lin, C.-H., & Kuo, B.-C. (2019, Jul). Application of the iAFM to computerized dynamic assessments. *Advanced Learning Theories, Technologies, Applications & Impacts 2019*, Memphis, TN.
- Yudelson, M. V., Koedinger, K. R., & Gordon, G. J. (2013). Individualized bayesian knowledge tracing models. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Artificial Intelligence in Education. AIED 2013. Lecture Notes in Computer Science* (vol 7926). Springer. [https://doi.org/10.1007/978-3-642-39112-5\\_18](https://doi.org/10.1007/978-3-642-39112-5_18)

測驗學刊，第 71 輯第 3 期