

## 壹、緒論

當研發電腦化適性測驗 (computerized adaptive testing, CAT) 時, 是否允許受試者回顧並修改之前的作答反應會是一個重要的爭論 (Bowles & Pommerich, 2001; Wise & Kingsbury, 2000)。而可允許的CAT則被稱為可回溯的電腦化適性測驗 (reviewable computerized adaptive testing, Reviewable CAT)。在一般CAT中, 是根據本次作答的答案正確與否當作下一道試題挑選的考量, 因此, 無法像一般的紙筆測驗可以跳題作答; 且考量受試者能力估計的準度和效度, 受試者一經作答, 將無法進行回溯的動作。所以, 大部分的CAT很少提供機會給受試者回顧試題及修改答案 (Mills & Stocking, 1996; Papanastasiou & Reckase, 2007; Stone & Lunz, 1994; Vicino & Moreno, 1997; Vispoel, 1998; Wise, 1996)。Vispoel (1998) 的研究指出, 有85%的受試者渴望有試題回溯的機會。更有甚者, 測驗焦慮指數愈高的受試者, 愈期待有回溯的機會 (Vispoel, 1998)。由大多數受試者的觀點來看, 在CAT上進行回溯, 是有益且必須的 (Vispoel, 2000)。然而, 受試者改變答案後, 隨之改變的作答反應 (response pattern) 可能會形成不合理的情況, 進而影響能力重新估計的結果。如何能增加受試者能力重估計的精確度? 成為待解決的問題。

此外, 目前多數研究可回溯之CAT大多建立在試題反應理論 (item response theory, IRT) 的基礎上, 但IRT主要是提供受試者整體潛在連續特徵的報告, 並不適合用來提供受試者是否精熟於某項認知概念的報告。而認知診斷模型 (cognitive diagnostic models, CDMs) 則可提供上述認知概念精熟之報告, 以便提供於補救教學之參考。是故, 可回溯之CAT是否適合用於CDMs, 便成為了待解決的問題。

因此, 為了讓CAT更貼近受試者之需求與測驗的合理性。且讓測驗的診斷報告能更符合補救教學之所需, 讓教師能更方便了解受試者不足之處。本研究欲探討在CDMs下, 結合作答回溯機制, 比較不同回溯法對能力估計的影響。

### 一、認知診斷模型

近年來診斷受試者的認知概念對教育評量的需求而言是與日俱增, 因此, Baxter與Glaser (1998) 及Nichols與Sugrue (1999) 認為, 在評量設計中應加強受試者的認知特徵。不同於用來診斷受試者廣義潛在特質的IRT, CDMs是一個以認知概念為主的診斷模型, 它主要在診斷受試者是否精熟某些認知概念。也將受試者在認知概念上的表現, 簡單地區分為精熟 (mastery) 與不精熟 (non-mastery) 的二元計分。認知概念組合則用一個向量 $\alpha_i = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{ik})$ 來表示第*i*位受試者是否

精熟第  $k$  個認知概念，其中  $\alpha_{ik} = 1$  代表精熟， $\alpha_{ik} = 0$  代表未精熟。而試題答對的機率則受到認知概念組合（pattern）的影響，進而用來估計受試者的概念精熟與否。因此，CDMs除了在教師對學生的個別化補救教學方面提供更詳細的協助訊息，也可提供自我的學習方向與目標給能力較佳的學生。

CDMs為了表示試題與認知概念之間的關聯，故需建立一個由學科專家所設計的  $Q$  矩陣（ $Q$ -matrix），用以界定答對試題所需要的認知概念（Tatsuoka, 1985）。假設有一份具  $J$  題試題及  $K$  個認知概念的試卷，則  $Q$  矩陣的維度為  $J \times K$ ，其定義如公式（1）所示：

$$q_{jk} = \begin{cases} 1 & \text{if item } j \text{ requires attribute } k \\ 0 & \text{else} \end{cases} \quad (1)$$

其中，

$q_{jk}$ ：表示要答對試題  $j$ ，是否需要具備認知概念  $k$ 。

對於CDMs而言，目前已有多種被開發且應用的模型。而其中由Junker與Sijtsma（2001）所提出的DINA模型（Deterministic Input, Noisy “And” Gate model）是屬於較基本且因其僅具有粗心及猜測兩試題參數，故為較簡單且易於說明之模型，再加上de la Torre與Douglas（2004）的研究顯示，DINA模型具有良好之模式適配度。因此，本研究使用DINA模型，以下介紹之。

## 二、DINA模型

DINA模型適用於二元計分的認知診斷測驗，也是在CDMs中最常被用來當作基礎評估的工具。此模型假設受試者是否能答對該試題，受到受試者是否具備該試題所需的認知概念，以及受試者是否受到猜測（guess）及粗心（slip）的影響。因此，在DINA模型中，受試者答對試題的機率如公式（2）所示：

$$P(X_{ij} = 1 | \alpha, s, g) = (1 - s_j)^{\eta_{ij}} g_j^{1 - \eta_{ij}} \quad (2)$$

其中，

$$\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{Q_{jk}}$$

$$s_j = P(X_{ij} = 0 | \eta_{ij} = 1)$$

$$g_j = P(X_{ij} = 1 | \eta_{ij} = 0)$$

$\eta_{ij}$ ：表示受試者是否具備試題所需的所有認知概念，完全具備為1；反之則為0。