

第一章 緒論

本章共分為七小節，包括研究背景與動機、研究目的、研究程序、基本假設、研究範圍、研究限制及論文架構，茲分別說明如下：

第一節 研究背景與動機

隨著電子化學習(E-Learning)環境技術的快速發展，有越來越多的學習者經由教學網站獲取知識，並透過教學網站內的測驗系統來了解自己的學習成效，因此線上學習及線上測驗已成為資訊教育領域的重要議題。然而，教育的目的是要引發學生行為的改變，因此教師必須善用測驗來評量學生學習後行為改變的情形，所以測驗在各種學科中，均扮演著重要的角色(林璟豐，民 90)。此外，就教學而言，測驗的主要目的在於增進學習的效果，所以在學習的各個階段以及擬定教學計畫時，必須將測驗納入成為整個教學過程中的一部份(陳英豪等，民 71)。因此，網路教學系統中教材內容及試題的配合，就成了網路教學系統中相當重要的工作。

測驗是評量學習成果的有效方法，傳統紙筆測驗，無論是出題、考試、閱卷、成績計算等作業，皆由人工進行處理，而受測者通常被指定在相同時間及地點進行考試，因此導致此方法的測驗效率不高，且人工作業出錯的機會亦相對較大(何榮桂,民 79;周文正,民 87)。由於電腦輔助測驗(Computer-Based Testing, CBT)具有提高測驗效率、減少測驗時間、降低測量誤差、提供多元豐富的試題呈現方式、獲得測驗相關的其他資訊等優勢，近年來國內外重要的測驗及檢定，已紛紛採用 CBT 的方式來施測及自動評分(何榮桂,民 79; 何榮桂,民 88; 周文正, 民 87)，例如國內的托福考試已採用電腦輔助測驗方式進行，而考選部也正積極規劃國家考試採行電腦化測驗作業，並期望能達成考試技術的重大突破(考試院,民 92)。

美國心理學會(American Psychological Associate, APA)在 1986 年，發表了以電腦為基礎的測驗在發展、使用以及解釋上的指導原則，此原則強調電腦測驗需與傳統紙筆測驗有同樣的效力(Bugbee,1996)。國內一些研究者也指出，現今實施線上電腦輔助測驗在教學活動中的確有其必要，而且使用電腦進行測驗時，其測驗效果應該與傳統紙筆測驗結果相同(何榮桂,民 89b;簡茂發,民 88)。然而，大部分的線上測驗系統，在設計上由於受到傳統紙筆測驗觀念的影響，因此考試題目的類型普遍仍多以單選題、複選題等認識型題為主，而忽略填充題等建構型題型的試題，因而造成某些認知層次的知識學習成效無法在測驗中被評量。雖然亦有部分的測驗系統提供填充題的測驗，但卻都只能做到答案的樣形(Pattern)比對，而並未針對答案進行語意分析的比對。林明達(民 87)也認為使用電腦批閱申論題及填充題的測驗題型並不容易，其中填充題的答案常有不同的寫法或有許多相同意義的詞彙，使得無法明確的指定標準答案。以計算機概論課程的填充題測驗為例，在傳統的紙筆測驗環境中，當受測者填寫答案「RAM」與题目的標準答案「隨機存取記憶體」在語意上相似或同義時，批閱者(教師)通常會根據受測者答案及標準答案間的語意相似程度斟酌給分，並非以電腦二元邏輯的樣形(Pattern)比對方式進行「非對即錯」的批閱給分。由於上述原因，目前提供填充題及申論題等開放式題型的電腦測驗系統，其評分結果與老師實際的評分結果可能並不相等，因此其測驗結果及評分效力也可能會受到質疑。由此可見，詞彙語意的辨識將是線上測驗系統發展需要努力的目標。

然而，在多位教師共用測驗系統的環境中，不同的教師對於填充題答案的詞彙語意及評分方式可能存在不同的認知。為了能實踐 Bugbee(1996)所提出「電腦測驗需能夠達到與傳統紙筆測驗具有同樣的效力」的觀念要求，因此延伸出了本研究的第一個主要的研究問題，也就是「如何在多位教師共用測驗系統的環境中，讓電腦具備教師個人化的評分風格與特質，並在包含填充題型的線上測驗中，電腦自動評分的結果能具有與紙筆測驗相同的評分效力？」

近年來 Bloom 等人(1956)的教育目標分類廣被國內外教育界採用，新版的 Bloom 認知領域教育目標被分成知識向度和認知歷程向度兩類，此理論可用以輔助教師發展有意義的教學規劃與測驗評量(葉連祺與林淑萍，民 92)。此外，在強調「自我學習」的線上學習環境中，學習者在知識向度及認知向度的學習成效，可藉由線上測驗來進行評量，若要達此目的，則線上測驗系統必須先了解每個試題所能評量的知識層次及認知層次。

一些研究指出，建立與維護測驗題庫是一項耗費人力的工作(蔡松齡，民 81a; 蔡松齡，民 81b; 林盈達、林義能，民 88)，當測驗系統題庫不足時，也容易造成題目出現頻率過高及學生直接記憶答案的情形。因此，如何提供豐富完整的題庫及有效率的出題方式，也成為遠距教學中的重要課題(何榮桂等人，民 85)。有研究指出，若是採用多位學者共同合作產製試題的策略來提升題庫產製效率，將容易產生題目重複或偏重某概念的情形(黃國禎等人，民 91)，試題若皆由老師的觀點來設計，則會無法兼顧學生的觀點與需求(林盈達、林義能，民 88)。然而，在大專院校中，並非每位專業系所的教師均瞭解教育理論及測驗理論的內容概念，因此大專院校教師在發展試題時，往往無法以 Bloom 的知識向度及認知向度概念來發展試題，因此試題即使發展完成後，教師也未必了解每個試題所能評量的知識類型及認知能力層次，且試題也未必符合試題編製的基本原則。綜合上述原因，延伸出了本研究的第二個主要的研究問題，也就是「如何讓電腦輔助教師產製具有 Bloom 分類概念且符合試題編製原則的試題？」

第二節 研究目的

基於前一節的研究動機，本研究的主要目的歸納如下：

- 一、以本體論、Bloom 教育目標分類理論、模糊理論、詞義關係、人工智慧為基礎，提出線上課程知識學習系統的建議架構，並以智慧型測驗評分及輔助教師產製具有 Bloom 分類概念題庫為研究核心，以期能降低教師製作測驗題庫及測驗評分的負擔。
- 二、探討影響線上測驗評分效力的因素來源，並提出解決方法，使電腦自動評分的結果能具有與紙筆測驗相同的評分效力。
- 三、在多位教師共享相似語詞庫的平台環境中，讓電腦能模擬教師個人的評分風格與特質，並藉由模糊理論來整合系統中所有教師的詞彙語意及評分認知，提供高品質的評分環境設定建議，降低教師建立語意詞彙及環境參數的負擔，並使系統提供之相似語意詞庫內容能與大多數教師的認知相近。
- 四、探討電腦輔助教師產製題庫的可行性及效益，並避免同一試卷中出現評量相同知識概念的重複試題。
- 五、結合教材知識本體及試題產製規則，產生包含 Bloom 教育目標分類概念資訊的是非題、單選題、複選題及多格填充題等試題，並用以評量學習者在知識及認知向度的

學習成效，強化試題在測驗中所產生的教育意涵。

第三節 研究程序

在 Yang(1995)的研究中指出，管理資訊系統領域的研究能夠區分成塑模(Modeling)、系統建置(Implementation)及實證研究(Empirical Studies)等三大類。本研究將先藉由先導實驗來瞭解研究問題，再以概念塑模(Conceptual Modeling)的方式，依據所關注的研究問題、提出的概念及探討之相關理論，提出一個系統架構，並經由系統設計建置一個雛形系統，最後再進行系統的成效評估。

本研究之研究程序可分為六個階段，分別敘述如下：

- 一、 評分效力問題來源先導研究：由於線上測驗使用電腦設備進行評分，因此可能會因為一些未知因素的干擾，導致線上測驗的評分結果與紙張測驗的評分結果不同，本研究將透過此階段的先導研究，以瞭解線上測驗可能導致評分效力降低的原因。
- 二、 教師發展測驗題庫的能力及問題先導研究：為了讓系統能輔助教師產製線上測驗之題庫內容，有必要先瞭解一般教師在人工出題時，實際的試題品質及發展困難，並探索教師人工出題所能涵蓋的試題類型、Bloom 知識類型及認知層次範圍，以挖掘出電腦在輔助教師產製試題時，可以提升教師發展試題效率及效能的機會。
- 三、 線上教材與試題結構知識本體概念設計：為了能有效的讓系統從線上教材中自動產生試題，本研究必須先瞭解教材講義及試題的本質結構及編排規則，因此本研究在進行初期必須確立出教材講義結構知識本體及試題結構知識本體內容。為使研究能以特定知識領域加以研究，本研究中將以「管理資訊系統」課程做為學習系統的教學課程應用範例。
- 四、 建立線上教材知識庫中的教材知識：本研究著重於利用教材知識本體結構所儲存的內容，輔助教師產製測驗題庫，因此本文將不針對教材知識經由斷詞、語意剖析、儲存至線上教材知識庫的技術方法進行討論，而是假設當線上教材知識均已依照教材知識本體結構儲存於線上教材知識庫中，再由電腦何利用本研究的本體結構及語意分析方法輔助教師產製測驗題庫。
- 五、 雛形系統發展：經由前三階段的研究與設計結果，規劃出能達成本研究目的之系統架構，並發展具備輔助教師產製題庫能力之線上課程知識學習系統雛形，雛形系統發展之步驟如下。
 - (一) 建立智慧評分機制：包括填充題模糊語意函數及推論規則、同音字相關字庫、全形半形相關字庫、英文大小寫處理模組、漏字模組的建立，以達成填充題型智慧評分的目的。
 - (二) 建立初始「相似語意關係詞庫」：當完成包含填充題題型之題庫建置完成後，需由專家(教師)針對每一個题目的填充題答案，依照該領域專業知識建立可相互共用之相似語意詞庫，以及非等同但近似的語意詞庫，並由專家建立「詞彙關係網路」，對於每一對相關的語意詞彙給予 0~1 間的相關係數，並藉此表示詞彙間的相似程度以利後續進行模糊評分(也就是系統可進行非全對或全錯的評分)。
 - (三) 建立語意詞庫的知識擴充機制：由於專家(教師)無法在測驗系統尚未啟用前，便能找出所有與填充題答案在語意上相同的全部詞彙，因此在正式啟用線上測驗系統後，學生如果發現系統無法辨識出相似語意的答案，可以透過「答案申訴

機制」來線上申訴。系統將會將此被提出申訴之題目及答案通知專家(老師)處理，並由專家(老師)決定是否該詞彙也應該加入至相似詞庫中，並同時決定該詞彙與標準答案的模糊關係程度。

- (四) 建立公式基模、知識樣式模型、詞彙語意庫：若要使系統能具備自動建立講義內容，必需先分析講義結構與知識結構的關係，並發展出各種構成講義內容的知識樣式模型。此外，中文語意變化莫測，若要能將知識做有效且正確的連結，則必須藉由中文語意詞庫來進行相似及相反詞彙的分析來達成。
- (五) 建立測驗樣式模型及測驗規則：若要使系統具備自動建立試題，必需先分析出所有測驗類型及試題結構，並探索試題與知識結構關係，以發展出各種構成試題的樣式模型與測驗規則。本階段所欲建立的測驗樣式模式，主要是根據二技、四技二專、高考、普考及勞委會各項技能檢定試題為基礎，進行題目內容之結構及規則分析，並從中萃取出本系統所需使用的測驗樣式模型及測驗規則。
- (六) 試題知識轉化：此階段在發展知識系統之試題產生模組，此模組將根據知識本體、知識關係、測驗模型、教材知識庫及試題產製規則等，由電腦輔助教師產製包含 Bloom 分類概念的線上測驗試題。此試題的題型則包括是非題、單選題、複選題及填充題類型，填充題則容許系統或教師在一個題目上設計最多 20 個空格答案，以增加題目的多元彈性。

六、系統成效評估：此階段將針對智慧評分機制是否能具備與紙筆相同的評分效力，以及輔助教師產製測驗題庫子系統的成效進行分析與評估。

第四節 基本假設

本研究的發展，包含了五項基本假設，分述如下：

- 一、教材文件的知識內容可由詞彙主題的觀點、詞義關係及特定文法規則來構成。
- 二、在多位教師共用的測驗平台環境中，不同的教師對於相同的詞彙語意會期望擁有相同的認知。
- 三、世界上的詞彙會不斷增生與修正，資訊系統所建立的詞彙資料亦需不斷的持續擴充與修正。
- 四、試題內容可歸納出特定的出題模式類型與規律，試題內容應來自教材文件中的範圍。
- 五、試題中評量知識向度及認知向度的資訊，可藉由特定的定義及結構來區別及辨識。

第五節 研究範圍

本研究的研究範圍可大略歸納成以下六點：

- 一、本研究的系統架構，雖包含線上課程、線上測驗、虛擬教室…等功能，但本研究僅針對電腦輔助教師產製測驗題庫，以及填充題智慧評分的方法及效益進行研究。
- 二、受限於研究目的與時間，本研究無法進行中文斷詞的方法技術發展，因此直接採用中央研究院 CKIP 中文斷詞系統來進行文章內容斷詞。
- 三、本研究在試題產製及線上測驗評分效力的議題中，僅以電腦可以自動評分的是非題、單選題、複選題及填充題等四類題型為主，不包含問答題、多媒體類型等試題

的發展議題。

- 四、本研究之線上測驗系統並未針對多媒體影音圖片的試題進行研究與探討，對於不同檔案格式的試題匯入與轉換功能，本研究因考量研究目的，亦未納入研究中。
- 五、本研究的線上測驗功能著重於平時學習過程的課程學習評量，此測驗方式與近年試題反應理論(IRT)及電腦適性測驗(CAT)著重單一時間點的能力鑑定不同，因此未包含適性測驗方法之研究。
- 六、由於本研究主要在提出一個能輔助教師產製測驗題庫及進行智慧評分的線上課程知識學習系統建議架構，並進行雛形系統的設計，因此本研究不包含系統架構對於學習者的學習成效、滿意度、學習行為及其他教育理論相關變數之實證研究。

第六節 研究限制

本研究的研究限制可大略歸納成以下六點：

- 一、由於考量課程教材取得的便利性，本研究僅選定「管理資訊系統」做為本研究教材內容及產製試題的應用領域。
- 二、教師雖然可能專精於某個課程領域，但由於中文文法及中文語意詞彙變化多元，管理者及教師難以在系統運作初期便建立完成所有專業領域中的中文知識結構及相似語意詞彙內容，也難以將所有的中文知識文章內或試題內的文法結構及規則在初期便能完整建立。因此，本研究透過語意詞彙、知識樣式管理、測驗樣式管理及測驗申訴等功能，來協助教師及管理者將無法辨識的語意詞彙及新規則擴充至系統中，藉著學生學習過程的申訴、教師教學經驗的累積，將可使本系統所能瞭解的詞彙、詞彙關係及文法規則變得更豐富。
- 三、由於中文輸入方法眾多，本研究僅以「注音輸入法」進行輸入法的智慧評分效力探究，而無法涵蓋所有輸入法。
- 四、本研究中電腦輔助教師產製的試題，僅能產生具備特定語句結構或語意關係的試題，具備情境性的靈活思考試題則仍需藉由教師以人工方式建立。因此，為了確保試題具備高品質與鑑別力，本研究另外提供教師可藉由人工方式來建立具備情境性的靈活思考試題。
- 五、由於研究時間的限制，本研究中的系統所產製出的大量試題，尚無法在本研究中針對每個試題進行試題鑑別度及難易度的分析與實證。
- 六、在委託教師進行人工產製試題及評估產製試題成效的實證研究中，由於教師均需投入2個月時間參與，並實際產製試題，因此無法進行大樣本的實證，僅能針對15位教師進行問卷、實驗及訪談等活動，以取得系統的成效分析結果。

第七節 論文結構

本論文之內容結構共分為六個章節，分述如下：

第一章「緒論」，主要針對研究背景與動機、研究目的、研究程序、基本假設、研究範圍、研究限制進行描述。

第二章「文獻探討」，主要是以研究目的為出發點，對於本研究相關之文獻加以整理與探討，以作為系統架構設計概念及擬定研究變數的依據，最後將綜合上述文獻結果而

作為本研究之基礎。

第三章「雛形系統設計」，主要針對評分效力問題來源先導研究、教師發展測驗題庫能力及問題之先導研究、教材與試題結構知識本體概念設計、系統課程知識學習系統架構設計、系統架構細部元件運作原理、電腦產製試題方法、評分機制設計原理、系統成效評估方式等內容進行描述與說明。

第四章「雛形系統建置」，主要針對本研究的系統環境、使用者觀點之系統整體功能說明、系統功能說明進行描述。在系統功能說明部分則區分為教師與管理者情境及學生情境兩個角度來探討系統功能。

第五章「雛形系統成效初步評估」，主要針對本研究之「輔助教師產製測驗題庫子系統」及「線上測驗與智慧評分子系統」進行系統成效評估，每一個部分除了針對實證研究的研究設計及資料分析結果進行說明外，也與現有測驗系統或教師人工處理的方式，進行功能效益的比較。

第六章「結論與建議」，主要針對本研究的研究成果、未來研究方向與建議及所能產生的貢獻進行描述。